REVIEW OF ACTUARIAL APPLICATIONS OF FUZZY SET THEORY

by

Y H YAKOUBOV and S HABERMAN

Actuarial Research Paper No. 105

Department of Actuarial Science and Statistics
City University
London

February 1998

ISBN 1 901615 23 5

ABSTRACT

This paper reviews the applications of fuzzy set theory to actuarial problems. Fuzzy sets are used to describe uncertain statements, where the uncertainty is due to the nature of the phenomenon, its perception by humans or arising from its complexity. The basic definitions and principles of fuzzy set theory are presented and fuzzy techniques, such as fuzzy numbers, fuzzy zooming of cash flows, fuzzy clustering, fuzzy expert systems and fuzzy decision making , which have been applied to actuarial and insurance problems are investigated. The areas of applications of the theory include financial mathematics, underwriting and risk classification, pricing of general insurance business, asset allocation, assets and liabilities matching and marketing. One of the key conclusions is that fuzzy set theory provides a promising way of treating uncertainty which is inherent to many actuarial applications and it would be a useful addition to the modelling tools used by actuaries.

KEYWORDS: fuzzy set theory, actuarial science.

CONTENTS

Introduction.

Mathematics of fuzzy set theory.
- Definitions.
- Fuzzy operations.
- Special fuzzy sets.
- Other useful concepts.

Fuzzy mathematics of finance.
- Introduction.
- Fuzzy arithmetics.
- Example.
- A life insurance application.
- Fuzzy 'zooming' of cash flows.
- Intervals of possibilities.
- Other applications.
- Conclusions.

Fuzzy pattern recognition in actuarial science.
- Pattern recognition.
- Fuzzy clustering.
- Actuarial applications.
- Conclusions.
- Appendix.

Fuzzy expert systems and fuzzy logic control.
- Theory.
- Actuarial applications.
- Conclusions.

Decision making in fuzzy environments.
- Decision making theory.
- Actuarial applications.
- Conclusions.

Fuzzy techniques as part of 'hybrid' models
- Introduction.
- Examples of ' hybrid' systems.
- Conclusions.
- Appendix.

Conclusions and further research.

1. INTRODUCTION

1.1 Fuzzy sets were introduced in 1965, when Lotfi A. Zadeh published his historic article 'Fuzzy sets', where he described the concept of a fuzzy set and some basic principles of fuzzy set theory(FST). The notion of a fuzzy set is very simple but has profound implications. A fuzzy set is a generalisation of the classical, Boolean set, for which an element either belongs or does not belong to the set. The fuzzy set is a set with ill- defined and imprecise boundaries. That is, the membership is a matter of degree, which is characterised by an appropriate membership function.

The fuzziness can be found anywhere in the real world. Concepts like 'beautiful', 'tall', 'young', 'satisfied customer' or statements such as 'our company has a high loss ratio' are fuzzy. A major concern in modelling the real world is accommodating and treating such vague and imprecise information. The opinion that vagueness/fuzziness is unscientific and thus should be avoided can no longer be justified.

1.2 Fuzzy set theory is about modelling uncertainty. Randomness is an important constituent of uncertainty. But does the notion of probability exhaust our notions of uncertainty? Does FST bring some new insights to the world of uncertainty or it is just a different interpretation and presentation of what we already know from the probability theory?

A classic review of the varieties of the notions of uncertainty, such as vagueness, nonspecificity and ambiguity is given by Black(1937). We can regard uncertainty as consisting of vagueness/ ambiguity/ fuzziness and randomness. Ostaszewski (1993, p. 2) states:

> ' Vagueness is associated with the difficulty of making sharp or precise distinctions among the objects studied. Ambiguity(*or in our terminology randomness*) is caused by situations where the choice between two or more alternatives is unspecified.'

Randomness describes the uncertainty of event occurrence, whether an event occurs or not and fuzziness relates to the degree to which an event occurs, not whether it occurs. Fuzziness refers to deterministic uncertainty. An anecdotal illustration is given by Kosko (1990): if there is 50% chance of finding an apple in the fridge, this is a state of affairs arrived at using probabilistic inference and if we consider that there is half an apple in the fridge, this would be another event. The two events are equivalent in terms of their numerical uncertainty, but one of them is random, the other is fuzzy.

Probability theory- the Kolmogorov (or frequentist) or Bayesian approaches- is successfully used to model the 'randomness' side of events and the fuzziness/imprecision is usually considered negligible: this is often the case in the physical sciences. But when we are confronted with the imprecision of natural language and of human perception, vagueness appears as a very important factor in modelling such phenomena.

There are many similarities between the two theories - FST and Probability Theory. Both describe uncertainty in a numerical manner, using numbers from the unit interval [0,1] and both theories involve the combination of their basic notions associatively, commutatively and distributively.
There are also differences. An interesting approach to them is given by Kosko (1990). He states that fuzziness occurs when, and only when, the first law of Aristotle's 'laws of thought' of non-contradiction (i.e. $A \cap A^c = \varnothing$) is violated, others being the law of excluded middle (i.e. $A \cup A^c = X$) and the law of identity (i.e. $A=A$), where A is a 'thing' and A^c is its opposite and X is the set of all 'things'. He writes:

> 'Classical logic and set theory assume that the law of noncontradiction and equivalently the law of excluded middle, is never violated. That is what makes the classical theory black and white. Fuzziness begins where Western logic ends.'

Probability theory is based on the theory of measure, while FST is not.
Increasing information about a phenomenon reduces the importance of probability and randomness, while if all facts are presented , fuzziness very often remains, i.e. a large hill is only roughly a mountain or a person with light injuries is only to some extent disabled.
Lemaire adds (1990) :

> 'Probability concepts are derived from considerations about uncertainty of propositions about the real world. Fuzzy concepts are closely related to the multivalued logic treatments of issues of imprecision in the definition of entitles. Hence, fuzzy set theory provides a better

framework than probability theory for modelling problems that have some inherent imprecision. ...

Classical probability theory has its effectiveness limited when dealing with problems in which some of the principal sources of uncertainty are non-statistical in nature.'

1.3 The first application of FST to insurance and actuarial problems is due to DeWit (1982). Some eight years later, fuzzy sets were rediscovered by Lemaire (1990), where the examples presented cover such areas as underwriting, reinsurance and financial mathematics. Under the auspices of the Society of Actuaries, Ostasziewski (1993) presented a review of possible applications of FST in actuarial science in areas such as the economics of risk, the time value of money, individual models, collective models and risk classification. Cummins and Derrig (1993, 1997), Derrig and Ostaszewski (1994, 1995) and Young (1996) have applied the fuzzy approach in non-life insurance context. Buehlmann and Berliner (1992), Berliner and Buehlmann (1993) have developed the theory of fuzzy 'zooming' of cash flows, while Babad and Berliner (1994, 1995) have used a slightly different approach towards the uncertainty- the Intervals of Possibilities. Other areas of applications have included asset allocation (Guo and Huang, 1996), underwriting and marketing (Hellman, 1995; Young, 1993), the matching of assets and liabilities (Chang and Wang, 1995).

1.4 In the following sections are presented the basic definitions and principles of FST and the fuzzy tools and techniques that have been applied to actuarial and insurance problems. Chapter 2 gives the basic definitions and principles, which represent in a rigorous, mathematical way the foundations of FST. Chapter 3 deals with the notion of a fuzzy number and its use in the time value of money, in calculating cash flows where amounts, interest rates, time of payments are uncertain quantities and in the theory of 'fuzzy zooming', as developed by Berliner and Buehlmann. Chapter 4 looks at fuzzy pattern recognition and gives examples of the applications of the fuzzy c-means clustering algorithm in actuarial modelling. Fuzzy approximate reasoning, fuzzy expert systems and fuzzy control systems are explained in Chapter 5 and applications in underwriting and risk classification are presented. Decision making in a fuzzy environment and its potential in the insurance and actuarial context are discussed in

Chapter 6. Chapter 7 presents a brief consideration to the so called 'hybrid' systems and their merits in modelling real world problems and presents the description of three such systems. Finally, we present our conclusions and some thoughts about further research in Chapter 8.

2. MATHEMATICS OF FUZZY SET THEORY

2.1 Definitions.

2.1.1 Given a universe of discourse U, **a fuzzy set A** in it is defined by

$$\mu_A : U \rightarrow M$$

μ_A is called the **membership function** of A and for each element x of U, gives the degree of membership of x in A, M is an ordered set, normally is taken to be the unit interval, $M=[0, 1]$. The definition is a generalization of the characteristic function of an ordinary set, where the characteristic function χ of a set A is

$$\chi(x) = \begin{cases} 0 & \text{if } x \notin A \\ 1 & \text{if } x \in A. \end{cases}$$

The two extreme values 0 and 1 represent, respectively, the lowest and highest degree of membership. The degree of membership could be interpreted as the "truth value" of a statement such as "x is a member of A".

2.1.2 The α - **cut** of a fuzzy set A is the crisp set A_α defined as

$$A_\alpha = \{ x \in U : \mu_A(x) \geq \alpha \}.$$

It could be thought of as an error interval whose truth value is α, i.e. there is a $\alpha\%$ belief that a particular element is in the set. For example a 0.90-cut of A contains all elements that are at least 90% in the set, and a 1-cut contains the elements that are for certain in the set. In the decision making process, the fuzzy decision, represented by a fuzzy set is replaced by an appropriate α - cut, and the so defined crisp set gives the set of optimal decisions.

The α - **level** in a fuzzy set A is

$$\Lambda_\alpha = \{ x \in U : \mu_A(x) = \alpha \}$$

and $$A_\alpha = \bigcup_{k \geq \alpha} \Lambda_k .$$

2.1.3 A fuzzy set A is **normal** if there is an $x \in U : \mu_A(x) = 1$.

If it is not and $m = sup\ \mu_A(x),\ x \in U$, then the new set $A' = A / m$, with $\mu_{A'}(x) = \mu_A(x)/m$ is normal.

2.1.4 $A \subset R$ is **convex** if for $c \in [0,1]$ and for any x and y

$$\mu_A(cx+(1-c)y) \geq min(\ \mu_A(x),\ \mu_A(y)\ )\ .$$

2.1.5 The **union** of fuzzy sets A and B is the fuzzy set $A \cup B$ with membership function

$$\mu_{A \cup B}(x) = max\ (\mu_A(x)\ ,\ \mu_B\ (x)),\ \ x \in U$$

This is the standard definition that includes the crisp case.

Other definitions that satisfy some particular set of axioms could be given. Ostaszewski (1993, pp.13-15) lists the axioms and shows the relationship of the generalised union (and intersection) with the so called *T- norms* and *T- conorms*.

The **intersection** of A and B is $A \cap B$ with membership function

$$\mu_{A \cap B}(x) = min\ (\mu_A(x)\ ,\ \mu_B\ (x)),\ x \in U.$$

The **complement** $-A$ of a fuzzy set A is defined as

$$\mu_{-A}(x) = 1 - \mu_A(x).$$

These operations are closely related to the connectives '*or*', '*and*' and *negation* respectively and naturally extend the corresponding definitions from standard set theory.

2.1.6 As an illustration, let

$$U = \{0,1,2,3,4,5,6,7,8,9,10\}$$

represents the possible outcome from an examination, rounded to the nearest multiple of 10, (i.e. 4 is assigned to 44 and 5 is assigned to 45). Let A and B denote the fuzzy sets that represent a student's expectation from two exams.

$A=\{(0,0),\ (1,0),\ (2,0.3),\ (3,0.9),\ (4,1),\ (5,0.9),\ (6,0.8),\ (7,0.5),\ (8,0.1),\ (9,0),\ (10,0)\}$

$B=\{(0,0),\ (1,0),\ (2,0),\ (3,0),\ (4,0.1),\ (5,0.2),\ (6,0.9),\ (7,1),\ (8,0.9),\ (9,0.2),\ (10,0.1)\}$

where $(2,0.3)$ denotes $\mu(2)=0.3$

then

$A \cup B=\{(0,0), (1,0), (2,0.3), (3,0.9), (4,1), (5,0.9), (6,0.9), (7,1), (8,0.9), (9,0.2), (10,0.1)\}$

$A \cap B=\{(0,0), (1,0), (2,0), (3,0), (4,0.1), (5,0.2), (6,0.8), (7,0.5), (8,0.1), (9,0), (10,0)\}$

$-A=\{(0,1), (1,1), (2,0.7), (3,0.1), (4,0), (5,0.1), (6,0.2), (7,0.5), (8,0.9), (9,1), (10,1)\}$.

2.1.7 In many applications, the minimum definition of intersection given above is too strict. Other definitions can be considered. Lemaire(1990) introduces the following three properties that the possible definitions of the intersection should satisfy:

Property 1: (*cumulative effects*):

$$\mu_{A \cap B}(x) < min\ (\mu_A(x)\ ,\ \mu_B\ (x)),$$

if $\mu_A(x)\ ,\ \mu_B\ (x)$ *are* <1.

The effect of two factors is worse than each separately.

Property 2: (*interactions between criteria*):

The effects of A and B on $A \cap B$ are not independent,i.e. the effect of a change of $\mu_A(x)$ on $\mu_{A \cap B}(x)$ may also depend on $\mu_B(x)$.

Property 3: (*compensation between criteria*)

An increase in $\mu_A(x)$ could be eliminated by a decrease in $\mu_B(x)$.

The minimum operator does not satisfy any of these properties.

2.1.8 The following definitions satisfy some or all of the three properties in 2.1.7

algebraic product *AB*, defined by

$$\mu_{AB}(x) = \mu_A(x)\ \mu_B(x),\ \ x \in U.$$

It satisfies all of the three properties.

bounded difference *A-B,* defined by

$$\mu_{A-B}(x) = max(0,\ \mu_A(x) + \mu_B(x) - 1),\ \ x \in U.$$

It satisfies only the properties 1 and 3. Indeed the change in *A-B* caused by a change in $\mu_A(x)$ is independent of the value of $\mu_B(x)$ as long as $\mu_A(x) + \mu_B(x) - 1 > 0$.

Hamacher operator, *H,* which depends on *p* is defined by

$$H(A,B,p) = \frac{\mu_A(x)\mu_B(x)}{p + (1-p)[\mu_A(x) + \mu_B(x) - \mu_A(x)\mu_B(x)]} \qquad 0 \le p \le 1$$

when $p=1$ the Hamacher operator reduces to the algebraic product, which is the intersection that allows maximum interaction. In general the following inequalities are in force

$$\mu_{AB} \le \mu_H^p \le \mu_H^q \le \mu_{A \cap B} \qquad q \le p\,.$$

The degree of interaction decreases when p decreases. The Hamacher operator satisfies the three properties in 2.1.7.

Yager operator, *Y*, depends on *p* and is defined by

$$Y(A,B,p) = 1 - min\ \{\ 1,\ [\ (1- \mu_A(x))^p\ + (\ 1 -\ \mu_B(x))^p]^{1/p}\},\ \ p\ \ge 1$$

When $p=1$ the Yager operator reduces to the bounded difference and when $p \to \infty$ it is equivalent to the minimum intersection. *Y(A,B,p)* is an increasing function of p. For $p>1$ it satisfies the three properties in 2.1.7.

The Hamacher definition falls somewhere between the minimum and the algebraic, and the Yager definition falls between the minimum and the bounded difference in modelling the possible interactions between the sets; the choice of p gives greater flexibility in determining the level of the interaction.

2.1.9 Let us consider the fuzzy sets defined in 2.1.6. The overall exam expectation could be given by the intersection of *A* and *B*.

The results of the different definitions are presented in the table below.

X	A	B	A∩B	AB	A-B	H(A,B,0.5)	Y(A,B,2)
0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0
2	0.3	0	0	0	0	0	0
3	0.9	0	0	0	0	0	0
4	1	0.1	0.1	0.1	0.1	0.1	0.1
5	0.9	0.2	0.2	0.18	0.1	0.19	0.19
6	0.8	0.9	0.8	0.72	0.7	0.73	0.78
7	0.5	1	0.5	0.5	0.5	0.5	0.5
8	0.1	0.9	0.1	0.09	0	0.09	0.09
9	0	0.2	0	0	0	0	0
10	0	0.1	0	0	0	0	0

From the results from the above table, we can see that the highest level of interaction is presented by the algebraic product, while the minimum intersection presents the least interaction. The Hamacher and Yager operators lie between these two extremes.

2.2 Fuzzy operations

Each fuzzy set can be manipulated by the fuzzy operations of concentration, dilation or intensification. These are defined as follows.

2.2.1 **Concentration** of *A* is a fuzzy set *CON(A,a)* defined by

$$\mu_{CON(A,a)}(x) = \{ \mu_A(x) \}^a,\ a>1.$$

The degree of membership of each element that belongs partially to A is reduced. The less the element is in the set, the more is its membership value reduced.

2.2.2 The opposite operation is **dilation,** *DIL(A, a)*

$$\mu_{DIL(A,\ a)}(x) = \{ \mu_A(x) \}^a,\ 0<a<1.$$

The degree of membership of each element that belongs partially to A is increased.

2.2.3 **Intensification** of A is the fuzzy set *INT(A)*, defined by

$$\mu_{INT(A)}(x) = \begin{cases} 2\mu_A^2(x) & 0 \le \mu_A(x) \le 0.5 \\ 1-2(1-\mu_A(x)^2 & 0.5 < \mu_A(x) \le 1 \end{cases}$$

The degree of membership of elements that are at least half in the set is increased and the degree of membership of those which are not is decreased.

These operations could be used in constructing fuzzy expert systems or in decision making procedures. They enable us to weight the significance of the different factors or groups of factors that influence the outcome of the model.

2.3 Special fuzzy sets.

2.3.1 **A fuzzy number** is a fuzzy set $A \subset R$, characterized by four real numbers a_1, a_2, a_3, a_4 and defined as(Lemaire 1990):

(a) $\mu_A(x)$ is continuous

(b) $\mu_A(x) = 0$ for $x \in (-\infty, a_1]$ and $[a_4, +\infty)$

(c) $\mu_A(x)$ is strictly increasing in $[a_1, a_2]$ and strictly decreasing in $[a_3, a_4]$

(d) $\mu_A(x) = 1 \quad x \in [a_2, a_3]$

This is represented in Figure 2.1.

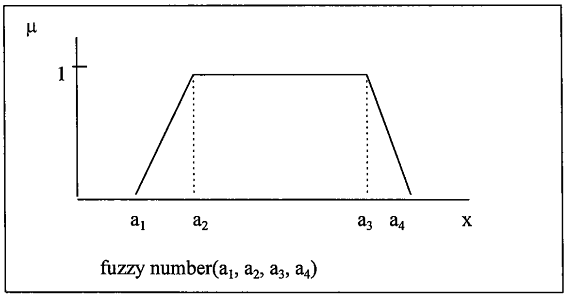


Fig.2.1 Fuzzy number

A more general definition of a fuzzy number is given by Ostaszewski(1993, p.11). It is defined as a normal, convex fuzzy set, with membership function that is continuous and vanishes outside an interval *[a,b]* of the real line.

A fuzzy number can be interpreted as a fuzzy subset of the real line whose highest membership values are clustered around a given real number or a real interval.

If $a_1 = a_2 = a_3 = a_4$ then A is an ordinary number.

We can say that A is a positive number if $a_1>0$ and a negative number if $a_4 < 0$.

2.3.2 A **triangular fuzzy number (TFN)** A is a fuzzy number with linear increasing and decreasing parts of the membership function and $a_2 = a_3$, as depicted in Figure 2.2.

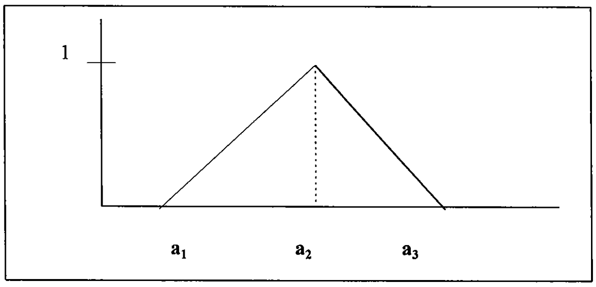


Fig.2.2 Triangular Fuzzy Number

2.4 Other useful concepts.

2.4.1 It is possible to define fuzzy analogues to many other elements from the fields of algebra, analysis and other branches of contemporary mathematics. For our purposes, we will present two definitions of fuzzy inequalities. Further discussions of this and similar extensions can be found in Zimmerman (1991), Dubois and Prade (1980) or in the specialised journals (e.g. 'Fuzzy sets and systems'). For example, another useful concept is that of fuzzy equations (Buckley, 1990; Cummins and Derrig, 1997).

2.4.2 If f is a conventional function (i.e. a mapping to the set of real numbers, $f: X \to R^1$), a fuzzy inequality is denoted as

$$f(x) \succ b\,(\theta)$$

and the meaning is that $f(x)$ is 'somewhat' bigger than b and the degree of being 'somewhat' bigger is controlled by the *tolerance* θ, $0 \leq \theta \leq 1$ (Zimmerman, 1991).

The solution is the fuzzy set A with a membership function μ_A,

$$\mu_A(x) = sup\{y \mid f(x) \geq b - \theta(1-y)\}$$ and y is a number between 0 and 1.

The above defined solution comprises the solution under the conventional approach with certainty and it also includes small intervals around the border points to some degree.

2.4.3 A second approach to fuzzy inequalities uses a fuzzy mapping and in fact, consists of comparing a fuzzy number with a 'crisp' number.

A *fuzzy mapping* is a function which returns as values (triangular) fuzzy numbers, $f: X \to TFN(R^1)$, where $TFN(R)$ is the set of all (triangular) fuzzy numbers.

The solution is the fuzzy set B, with a membership function μ_B,

$$\mu_B(x) = sup\{y \mid v_{x-}(1-y) \geq b\},$$

where v_{x-} is the inverse of the decreasing part of fuzzy number's $f(x)$ membership function (see section 3.2.1) and y is a number between 0 and 1.

2.4.3 Other definitions of fuzzy inequalities are possible. For example, Chang and Wang (1995) consider two further examples, based on the above definitions and the same authors provide some elements of fuzzy calculus as well.

3. FUZZY MATHEMATICS OF FINANCE

3.1 Introduction

The elements that characterise many financial problems are vague and imprecise. For a given cash flow, when calculating the present value, the future rate of interest, time of payment or amount of each payment are not known and appropriate estimates are needed. When there are sufficient statistical data and confidence that the future development of the cash flow will follow the past experience, stochastic models are used. Fuzzy numbers could be used when the whole picture is not clear and the available information is scarce or it is not possible to make a definite decision because of the subjective nature of the human perception. The fuzzy numbers are used when the complexity of models makes it difficult to obtain practical results. Then methods based on fuzzy arithmetics could be used in order to simplify the original model.

3.2 Fuzzy arithmetics

3.2.1 Operations on fuzzy numbers, such as summation, product and power are defined, using the Extension Principle of Zadeh. It is a method of extending functions to fuzzy sets (see Ostaszewski, 1993, pp.11, 28) and in an elementary form it was presented by Zadeh (1965) and developed by Dubois and Prade (1980).

3.2.1 **The sum** of two fuzzy numbers A and B

$$S = A + B$$

is defined as the fuzzy set with the following membership function

$$\mu_S(s) = \max_{s=x+y} \min(\mu_A(x), \mu_B(y))$$

It can be proved that the so defined fuzzy set is a fuzzy number and that the summation is (in this context) an associative and commutative operator (Dubois and Prade, 1980). This is illustrated in Figure 3.1.

The properties below follow from this definition.

1. $\mu_S(s) = 0$ $\quad$ *for* $x \in (-\infty, a_1+b_1) \cup (a_4+b_4, \infty)$

2. $\mu_S(s) = 1$ *for* $x \in (a_2+b_2, a_3+b_3)$

3. if μ_+ and μ_- are the increasing and decreasing parts of a fuzzy number's membership function and $v_+=[\mu_+]^{-1}$, $v_-=[\mu_-]^{-1}$ are the inverse functions then

$$v_{S+} = v_{A+} + v_{B+}$$

$$v_{S-} = v_{A-} + v_{B-}$$

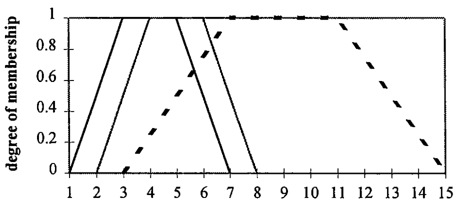


Fig. 3.1 Sum of two fuzzy numbers. A=(1, 3, 5, 7), B=(2, 4, 6, 8); A+B=(3, 7, 11, 15).

3.2.2 A similar definition, based on the Zadeh's Extension Principle, can be given to **the product** P of two fuzzy numbers A and B, P = A⊗B, which is defined as

$$\mu_P(p) = \max_{p=xy} \min(\mu_A(x), \mu_B(y))$$

It can be shown that P is a fuzzy number and is characterized by

$$p_1=a_1b_1,\ p_2=a_2b_2,\ p_3=a_3b_3,\ p_4=a_4b_4$$

and has the following inverse parts of the membership function:

$$v_{P+} = v_{A+} * v_{B+}$$

$$v_{P-} = v_{A-} * v_{B-}$$

Also it can be proved that the product is associative, commutative and distributive with respect to summation (Dubois and Prade, 1980).

3.2.3   The **nth power** of A is defined recursively as

$$A^n = A \otimes A^{n-1}$$

3.3 Example

3.3.1 We will calculate the net single premium of a £100 10-year pure endowment policy on a life aged 25.

3.3.1.1 We assume that the rate of interest i (assumed to be constant) for the next 10 years is a fuzzy number, clustered around the 7% level so that the fuzzy number $1+ i$ has the following membership function:

$$\mu_{1+i} = \begin{cases} 100x - 105 & 1.05 \le x < 1.06 \\ 1 & 1.06 \le x < 1.075 \\ 44 - 40x & 1.075 \le x \le 1.1 \end{cases}$$

which is shown in Figure 3.2.

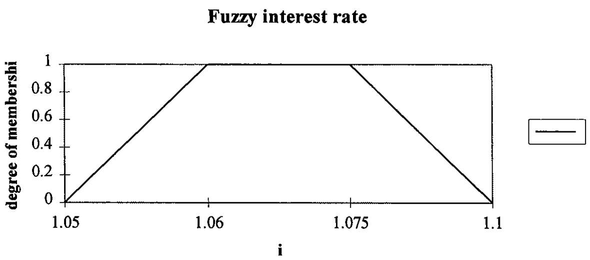


Fig. 3.2 Interest rate factor *1+i=(1.05, 1.06, 1.075, 1.1)*

If the inverse functions of the monotonic parts are v_1 and v_2 then

$$\begin{cases} v_1(y) = 1.05 + 0.01y \\ v_2(y) = 1.1 - 0.025y \end{cases} \qquad y \in [0,1]$$

3.3.1.2 The net single premium (NSP) for the pure endowment is

NSP={sum assured}{prob. that payment is made}* {discount factor}*

The discount factor $DF = (1 + i)^{-10}$, where $(1 + i)$ is the fuzzy number defined above, and

$$\begin{cases} v_{DF+} = (1.1 - 0.025y)^{-10} \\ v_{DF-} = (1.05 + 0.01y)^{-10} \quad y \in [0,1] \end{cases}$$

and therefore DF is approximately the fuzzy number *DF= (0.3855, 0.4852, 0.5584, 0.6139)*

The survival probability (mortality table A1967/70) is

$${}_{10}p_{25} = l_{35} / l_{25} = 0.99$$

Then

$$NSP = 100*0.99*(1 + i)^{-10}$$

and using the definitions of product and n^{th} power, *NSP* is found to be approximately the fuzzy number NSP=(38.16, 48.03, 55.28, 60.78), shown in Figure 3.3.

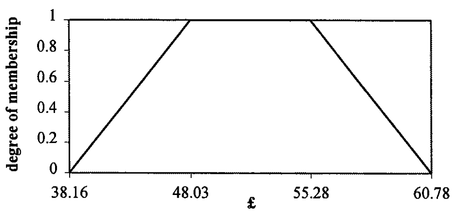


Figure 3.3 Net single premium as a fuzzy number, *NSP=(38.16, 48.03, 55.28, 60.78).*

3.3.1.3 Next it could be assumed that ${}_{10}p_{25}$ is also a fuzzy number. Specifically we assume that:

$$\mu_p(x) = \begin{cases} 0 & x \le 0.9 \\ \dfrac{100}{9}x - 10 & 0.9 < x \le 0.99 \\ 100 - 100x & 0.99 < x \le 1 \\ 0 & x > 1 \end{cases}$$

or using the inverse functions

$$\begin{cases} v_{p+} = 0.09y + 0.9 \\ v_{p-} = 1 - 0.01y & y \in [0,1] \end{cases}$$

And then the inverse functions for the fuzzy premium become

$$\begin{cases} v_{premium+}(y) = 100 * v_{p+} * v_{DF-} = 100(0.9 + 0.09)(1.05 + 0.01y)^{-10} \\ v_{premium-}(y) = 100 * v_{p-} * v_{DF-} = 100(1 - 0.01y)(1.1 - 0.025y)^{-10} \end{cases}$$

and the fuzzy number that represents the NSP is *(34.70, 48.03, 55.28, 61.39)*

The number is fuzzier because of the fuzzy survival probability that has been introduced.

It is also possible to 'fuzzify' the time of the payment (Buckley 1987) or the amount of payment (Buehlmann and Berliner, 1992).

3.3.2 We now revisit the same example but using a different approach. Assume that the rates of interest each year are represented by independent and identically distributed random variables i_t which have lognormal distribution with parameters μ and σ^2, i.e. $\log i_t \sim N(\mu, \sigma^2)$

3.3.2.1 The present value(PV) of 1 payable after 10 years equals to

$$PV = \{(1+ i_1)\,(1+i_2) \ldots (1+i_{10})\}^{-1}$$

$$\log PV = -\,\{\log(1+ i_1) + \log(1+i_2) + \ldots + \log(1+i_{10})\}$$

The right-hand side of the last equation is a sum of independent normally distributed random variables with mean $-\mu$ and variance σ^2. Then

$$logPV \sim N(-10\mu,\ 10\sigma^2).$$

3.3.2.2 If we take $E(1+i_t) = m = 1.07$ and $Var(1+i_t) = s^2 = 0.015^2$

and using the relationships between the parameters of the lognormal distribution and its mean and variance:

$$\begin{cases} m = \exp(\mu + \frac{\sigma^2}{2}) \\ s^2 = \exp(2\mu + \sigma^2)(\exp\sigma^2 - 1) \end{cases} \quad \textit{and the inverse} \quad \begin{cases} \mu = \log\frac{m}{\sqrt{1+(\frac{s}{m})^2}} \\ \sigma^2 = \log(1+(\frac{s}{m})^2) \end{cases}$$

we come to the following distribution for logPV

$$logPV \sim N(-.72146,\ 0.00346)$$

or $\quad mean(PV) = 0.5094$

$\quad s.d.\ (PV) = 0.0226$

then the $NSP = 100*0.99*PV$ is a random variable with *mean = 50.43* and *standard deviation = 2.24*

3.3.3 The result is similar to that calculated with the use of fuzzy numbers. In the fuzzy case, the length of the interval of most possible values is about £7. In the stochastic result, if we assume that a "confidence" interval for the NSP has length 3*s.d., then it is again about £7.

The use of the stochastic approach needs a solid theoretical basis, but allows implicitly for the inclusion of any statistical data available (estimating mean and variance from the past experience). The use of fuzzy numbers is simpler conceptually and allows greater flexibility (in determining the shape and actual values of the fuzzy numbers). The fuzzy approach is easily modified to include the cases when not only the interest rates, but also the amount or time of payment are not certain quantities. Both methods need an estimate of the future interest rates.

3.4 A life insurance application.

Using the same fuzzy approach, we can find the actuarial values of other benefits. Ostaszewski (1993) calculates the net single premium for a 2 year term assurance with the benefit payable at the end of year of death. The interest rates are fuzzy numbers and are read from a fuzzy yield curve. The yield curve is assumed at time 0 to be equal to the set of current short term interest rates, the 10 year rate is taken to be a triangular fuzzy number, having degree of membership of 1 at the current 10 year bond yield and the yield curve is assumed to increase linearly from the level of the short term interest rates to the fuzzy 10 year yield with increasing "fuzzification". The process of calculation is quite laborious but can be easily handled by computer.

The author suggests the use of a fuzzy approach in calculating future reserves and the possibility of showing at an early stage when the premiums charged appear low and a correction is needed.

It would be interesting to see how the fuzzy approach, with its flexibility in choosing the level and the form of the fuzzy yield curve compares to the more traditional deterministic modelling, scenario testing or stochastic modelling. An extension to the above example would be to perform a fuzzy scenario testing, with different assumptions regarding the fuzzy yield curve.

3.5 Fuzzy 'zooming' of cash flows.

Buehlmann& Berliner (1992) and Berliner& Buehlmann (1993) take the applications further by developing a theory that allows complex and uncertain cash flows to be replaced by some specific payments that approximately detect the sensitivity of the original cash flow to the interest rate movements.

3.5.1 In their approach we consider a given set of cash flows $\{G_t\}$ t = 1, 2, ..., n . Then

the *duration* D is defined as $D=\dfrac{\sum t v^t G_t}{PV}$

the *dispersion* M^2 is defined as $\quad M^2 = \frac{\sum t^2 v^t G_t}{PV} - D^2$

the *convexity* CO is defined as $\quad CO = v^2(M^2 + D^2 + D)$

and the *fuzzy payment equivalent to the cash flows* G_t is defined as the ordered pair

$$\{ PV(1+i)^D ;\ F[D-\sqrt{M^2}, D, D+\sqrt{M^2}]\}$$

where F is a triangular fuzzy number, PV is the present value of the payments $PV = \sum G_t v^t$ and

$v = \frac{1}{1+i}$.

3.5.2 The duration characterises the sensitivity of the cash flow to interest rate changes. The dispersion can be regarded as a measure of the deviation of a cash flow from a zero coupon bond of the same duration (Buehlmann& Berliner, 1992). An important feature of the concepts of duration and convexity (as defined above) is that the sum of the durations and the convexities of securities in a portfolio, weighted by their present values is equal to the duration and convexity of the portfolio as a whole.

Replacing the cash flow by an equivalent fuzzy payment reduces the complexity that arises from the pattern of the payments, but retains the main characteristics. The first element of the equivalent fuzzy payment $PV(1+i)^D$ gives the "value" of the cash flow at time point D, which could be thought of as the cash flows' "centre of gravity". The second element takes the duration as the central point of the triangular fuzzy number and the corner points give maximum "stretches" of the cash flow payments, because M^2 provides (as stated above) a measure of the deviation of the cash flow from a zero coupon bond.

3.5.3 The equivalent fuzzy payments are easily extended to *fuzzy zoomings* of a cash flow as follows:

- a fuzzy zooming of first order is the fuzzy payment equivalent to the cash flow.

- the original cash flow is divided into two cash flows by the duration D, i.e. one with payments between times 0 and D and second with payments between times D and n.

- for these two partial cash flows the durations and dispersions D and M are calculated, then the respective equivalent fuzzy payments are determined. These two payments represent the fuzzy zooming of second order.

- each of the two partial cash flows is divided again into another pair using the same approach, giving altogether four partial cash flows. Each one is assigned an equivalent fuzzy payment and these represent the fuzzy zooming of order four and so on...

- a fuzzy zooming of order k is represented by the k equivalent fuzzy payments $(PV_1(1+i)^{D}{}_1, F_1)$, $(PV_2(1+i)^{D}{}_2, F_2)$, ... , $(PV_k(1+i)^{D}{}_k, F_k)$.

3.5.4 Two cash flows react similarly to interest rate changes if they have identical fuzzy zoomings of some specified order. The equality of these zoomings could be used to obtain a result which is intermediate between a perfect hedge and an immunisation position. As stated in Buehlmann& Berliner (1992) the fuzzy zooming of first order is 'a much improved form of immunisation, because it takes into account the "stretching" by $\sqrt{M^2}$ and the fuzzy zoomings of higher order are nearly a perfect match'.

The fuzzy zooming could be generalised (Berliner& Buehlmann, 1993) by dividing the original cash flow into k disjunctive partial cash flows and for each one, an equivalent fuzzy payment is defined as above.

3.5.5 A simple illustration of a fuzzy zooming of a cash flow in life annuities is given by Buehlmann& Berliner (1992) and is reproduced below.

3.5.5.1 A deferred annuity financed by a single premium is issued to a life aged 20, the payments start at age 60 and the rate of payment is 1 p.a., and an annual rate of interest of 5% is assumed. For simplicity the cash flow paid by the life office is taken to be l_x (according to a mortality table, $l_0=100{,}000$, $x = 60, 61, \ldots$).

3.5.5.2 Then the following is calculated for the zooming of first order:

Duration	$D=48.1$ *(years)*
Present value	$PV=139\ 200$

Time value at time D $TV = 1\ 460\ 000$

Dispersion $M^2 = 46.9\ (years)^2$

PV and TV are expressed in monetary units.

Then the fuzzy zooming of first order is *(TV; TFN(D-M, D, D+M))* or

(1460000;TFN(41.25,48.1,54.95)) which is shown in Figure 3.4.

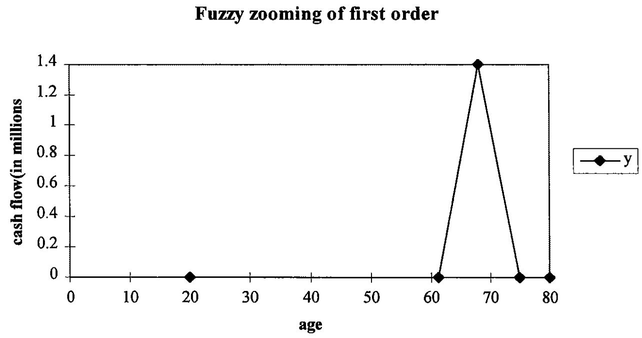


Figure 3.4 A fuzzy zooming of a cash flow in life annuities.

3.5.5.3 The fuzzy zooming of second order is calculated as mentioned above.

$D_{21} = 43.9$ *(years)* $D_{22}=55,8$ *(years)*

$PV_{21}=89\ 700$ $PV_{22}=49500$

$TV_{21}=765\ 000$ $TV_{22}=754\ 000$

$M_{21}^2 = 8\ (years)^2$ $M_{21}^2 = 25.9\ (years)^2$

And the ordered pairs

(765 000;TFN(41.0, 43.9, 46.7)), (754 000, TFN(50.7, 55.8, 60.8))

are the fuzzy zoomings of second order.

Thus a cash flow, consisting of many payments is replaced by an equivalent simpler cash flow, which is easier to analyse.

3.5.6 A second example (Berliner & Buehlmann, 1993) considers a geometrically increasing cash flow which is exposed to four different types of "interest rate shocks" and it is shown that generalised zooming of order five is a good approximation to the original cash flow, comparing its sensitivity to interest rate shocks.

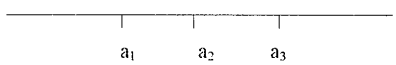
3.5.7 A potential area of use of this technique would be the modelling of assets and liabilities for a life insurer or a pension fund.

Given the simplicity of the fuzzy cash flows and the results of the fuzzy 'zooming' process, the above technique could replace the classical immunisation approach and given its alleged superiority, it could be used as an integral part of determining an appropriate matched position for an insurance company or a pension fund.

3.6 Intervals of possibilities.

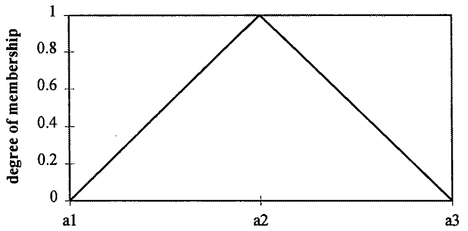
Babad and Berliner(1994, 1995) add a third approach to the treatment of vague and imprecise nature of the real world. Instead of using stochastic techniques or fuzzy set theory they introduce the *Interval of Possibilities* as a tool for modelling, analysis and decision making when the underlying information is scanty.

3.6.1 An interval of possibilities is a subset of the real line and is used to replace exact numbers, without making assumptions about any probability measure or degree of membership. It is defined by an ordered triplet (a_1, a_2, a_3); a_1 is the infimum value of the interval and it is interpreted as the most pessimistic value, a_3 is the supremum value, the most optimistic value and a_2 is the plausible value, the "average" person's view regarding uncertainty. The ordinary real numbers are a limiting case when $a_1 = a_2 = a_3$.



Arithmetic operations such as addition, subtraction, multiplication and division of intervals are defined in a natural manner.

An interval of possibilities could be easily "upgraded" to a fuzzy set and more precisely a triangular fuzzy number, by defining the membership function as shown below.



3.6.2 Applications of the above concept are considered by Babad and Berliner (1995) and include an interpretation of the sliding scale excess of loss (XL) reinsurance treaty as an Interval of Possibilities, the pricing of a fixed income security, the pricing of a callable bond- a security with an uncertain maturity date and the calculation of the pure premium for a life assurance policy with mortality rates represented by intervals of possibilities.

3.6.3 One of the biggest advantages of this theory, the delay in elimination of the uncertainty to the latest possible moment in the decision making process, also leads to a very large interval for the final result that sometimes makes the analysis a rather futile exercise. Babad & Berliner (1995) present a way of overcoming the problem by an "actualisation" of the intervals i.e. when a value from an interval is chosen it is consistently used in all places where the interval is presented. This results in a reduction of the size of decision intervals because some values become impossible.
The idea of actualisation is used in a Stop Loss Reinsurance example.

3.6.4 The concept of Intervals of Possibilities is simple and easy to understand. It could be used as an initial step in analysing processes which are not well determined. Because of the apparent low level of mathematical difficulty, the concept could be attractive to people who do not want to experience 'heavy' mathematics or could be useful for explaining results, obtained using other more 'strict' methods, to non-actuaries, for example, the trustees of an pension fund or the directors of an insurance company.

3.7 Other applications.

Many actuarial problems can be presented using fuzzy theory terminology and solutions derived using the tools of the fuzzy mathematics of finance. Two further examples appearing in the literature are briefly presented in the next paragraphs.

3.7.1 Derrig and Ostaszewski (1995) consider a situation where the use of the fuzzy numbers provides a simple and easy-to-apply tool. They look at the hedging of the tax liabilities of a general insurer and the uncertainty in the underwriting and investment parameters is modelled with fuzzy numbers. This leads to fuzzy quantities for the effective tax rate, the rate of return and the liability hedge. They conclude that the use of fuzzy numbers instead of exact numbers provides a more realistic view of the variability of the outcome quantities.

An alternative solution for this problem would be the use of random variables, instead of fuzzy numbers, and simulation techniques for obtaining the distribution for the outcome quantities. Although the two approaches are different in their nature, a common characteristic is the significant amount of subjectivity which is inherent in determining the input quantities. It is likely that the two models would lead to comparable results and superiority of one over the other would depend on the processes that are modelled and their characteristics.

3.7.2 In Chang and Wang (1995), fuzzy analogues of the classical immunisation theory and the matching of assets and liabilities are presented. The conventional theories are "translated" into fuzzy terminology and then some results from fuzzy algebra are used to solve the resulting problems.

The use of the fuzzy approach brings a new view to the problems themselves, provides an extension to the solutions by using fuzzy techniques and, more importantly, allows the setting of the problems to be closer to the real world and provides some flexibility in the interpretation of the solutions.

3.8 Conclusions

Fuzzy numbers provide an alternative to use of exact numbers for important actuarial quantities like interest and inflation rates, and amounts and times of payment. They are a convenient way to include easily, in a mathematical model, statements such as 'approximately between 5% and 8 %' or 'around £20, 000'.

The use of fuzzy numbers instead of probability theory is justified when insufficient statistical data are available or the underlying characteristics are vague and imprecise. Then fuzzy numbers allow their user to incorporate his/her opinion or feelings about the future behaviour of the unknown phenomenon.

When the problem itself is complex, a fuzzy approach could lead to a model that is simpler to solve but which still gives reliable results.

The theory of fuzzy numbers is relatively simple and that makes it easy to apply and the changes in the sizes or shapes of the fuzzy numbers can be given an intuitive explanation.

A drawback of the use of fuzzy numbers is the high level of subjectivity . There could exist several 'reasonable' assumptions which may lead to very different decisions.

Although the concept of a fuzzy number is theoretically simple and easy to use, producing meaningful results, which are an improvement on the results arising from conventional methods, requires skills and awareness of the underlying real world processes. In this sense, there is no 'standard' methodology, but each problem requires a unique approach.

4. FUZZY PATTERN RECOGNITION IN ACTUARIAL SCIENCE

4.1 Pattern recognition.

4.1.1 Pattern recognition is one of the dynamic branches of modern mathematics. In includes research in many areas of contemporary science - artificial intelligence, linguistic, biological and medical sciences. Bezdek(1991) defines pattern recognition as "a search for structure in data". Modern pattern recognition techniques can be divided into two classes: mathematical (primarily cluster analysis) and non - mathematical. The techniques of mathematical pattern recognition are more universal and are applicable to a variety of situations.

4.1.2 The most effective and easy way of finding a structure in data- when possible- is the 'eyeball' technique (e.g. medical diagnosis). Often the data are not easily interpretable and the need for more powerful and sophisticated search procedures arises. Statistical pattern recognition techniques have been developed to analyse multidimensional data. However, there are data sets which have a non-probabilistic nature and fuzzy set theory offers a way to analyse such data.

4.1.3 The entire pattern recognition process can be presented by Figure 4.1 (Zimmerman 1991):

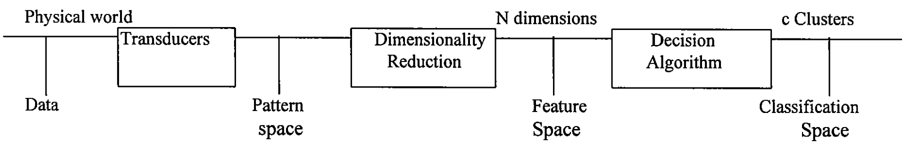


Fig. 4.1 Pattern recognition process.

The data are drawn from a real process; the form of the data could be quantitative, qualitative, numerical, pictorial, linguistic or any combinations of them. The data carry information about the process. Then the initial, raw data are transformed by the transducers to a format, which is suitable for analysis.

Finding a structure in the pattern space is defined by Bezdek (1981) as 'the manner in which this information can be organised so that the relationships between the variables in the process could be identified.' The feature space includes only those characteristics of the data that have a significant influence on the process. The classification space contains the clustering rules, implemented by the clustering algorithm.

4.1.4 In summary, the main issues for any pattern recognition problem are (Bezdek 1991, p.4):

- feature selection: the search for structure in the observed data.
- cluster analysis: the search for structure in data sets.

classification: the search for structure in data space or population.

4.2 Fuzzy clustering.

4.2.1 Let us denote a data set by X, i.e.

$X = \{x_1, x_2, \ldots, x_n\}$ where $x_i \in \Re^p$ are the observations and p is the number of features that characterise the process (already selected from the data space).

The aim of clustering is to divide X into c homogeneous subsets, called "clusters". The objects which belong to one cluster should be as similar as possible, and the objects from different clusters as dissimilar as possible.

To do this some parameters need to be determined:

1) c, the number of clusters, which is usually not known in advance

2) the clustering criterion, i.e. what is the explicit aim for the clustering. This is closely related to how we measure the similarity between the elements.

Mathematical properties of the data, e.g. distance, angle, curvature, connectivity, symmetry, intensity and so on would be used in the search for structure in the data.

4.2.2 It is important to bear in mind that there is not a clustering criterion that is universally acceptable and two reasonable criteria could exist for the data that lead to quite different results. As an illustration Bezdek (1981, p.45) provides an example, illustrated in Figure 4.2:

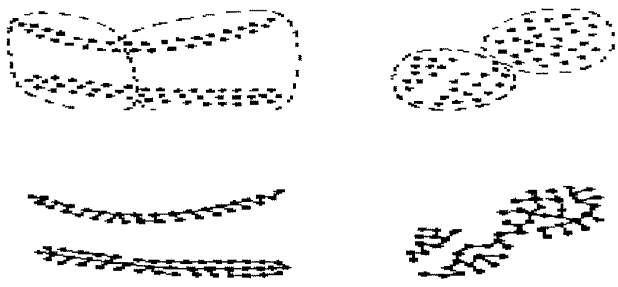


Fig.4.2 Clustering- success and failure.

The two data sets are divided into two clusters using:

1) a distance-based objective functional (Within Group Sums of Squares Criterion),

2) a distance-based graph theoretic method (Single-Linkage Criterion).

Obviously, the criterion that gives a good result in one of the data sets, performs badly in the other. Thus the success of a cluster analysis (crisp or fuzzy) depends on the investigator's ability to choose the criterion and the similarity measure that are most appropriate to the data.

4.2.3 The clustering algorithms are commonly categorised by their axiomatic basis ***(deterministic, stochastic, fuzzy)*** and then by the type of clustering criterion ***(hierarchical, graph-theoretic and objective functional criteria)***. Further subdivision can be made by considering the types of similarity measure. These terms are explained in more detail below:

deterministic basis refers to the fact that an outcome can, with absolute certainty, be predicted; for example, it has a certain functional form. In biology, for instance, the amount of bacteria at time t is often assumed to obey the law of exponential growth and the parameters of the functional dependence are uniquely determined. The uncertainty involved is called deterministic;

stochastic basis refers to random outcomes, which are unaffected by the environmental imprecision, but are inherent for the models. These models are called stochastic and are dealt with by statistical science.

In the first two cases, either the source of uncertainty is deterministic, but our ability to monitor it exactly is uncertain (deterministic), or the outcome itself is uncertain (stochastic). However ***fuzziness*** refers to uncertainty due to the subjective perception of the real world or to its complexity. For example, a statement such as ' this man is tall' is fuzzy.

Hierarchical methods create a hierarchy of partitions by successive merging or splitting of the clusters. In both cases, reallocation of one point at a time is considered, based on some similarity measure and the results are presented in the form of a tree-like structure, i.e. a hierarchy of nested clusters is generated. Deterministic and stochastic models are widely used, and a fuzzy hierarchical method is also available, in terms of "similarity trees", (defined in Zimmerman, 1991).

With ***graph-theoretic*** methods, the data set X is regarded as a set of nodes and the edge weights between pairs of nodes can be based on a measure of similarity between the nodes.
The criterion for clustering is some kind of measure of connectivity between the groups of nodes. The clustering strategy often involves the breaking of edges in order to form subgraphs.
Such methods are suited to special types of data, but they do not generate typical representatives of each subclass and therefore are useful for initial attempts at classification.
These methods are mainly deterministic and if the graph that represents the data is a fuzzy graph then different definitions of fuzzy connectivity lead to different clustering algorithms.

Objective functional methods allow the most precise formulation of the clustering criteria.

A local extremum of the objective functional is considered as an optimal solution to the clustering problem. Many different objectives are known (deterministic, stochastic and fuzzy), and an extensive coverage is given by Bezdek (1981).

4.2.4 The most frequently used clustering method (crisp and fuzzy versions) with applications in image recognition, medicine and many other areas of the contemporary science is the so called **c-means algorithm**. The classical, crisp algorithms generate partitions such that each element belongs to exactly one cluster. Often, however, it is difficult to find such distinct clusters and there exist elements that are "between" clusters. Then, the use of fuzzy clustering methods gives a more precise interpretation of the real data.

Let us consider the clustering method itself.

4.2.4.1 The data set is $X = (x_1, x_2, \ldots, x_n)$ and $x_i \in \Re^p$.

Each x_i is an observation and consists of p features:

$$x_i = (x_{i1}, x_{i2}, \ldots, x_{ip}),$$

where x_{ij} is the *j-th* feature of the observation x_i

The measure of dissimilarity between two elements x_k, x_l is taken to be their distance apart $d(x_k,x_l)$, where d is a suitably defined metric in X.

If the fuzzy clusters are

$$S_i \qquad i=1, \ldots, n$$

we can define

$$\mu_{ik} := \mu_{S_i}(x_k),$$

where $\mu_{S_i}(x_k)$ is the degree of membership of x_k in S_i.

The matrix

$$U = [\mu_{ik}] \quad i=1, \dots, c, \; k=1, \dots, n$$

is a *fuzzy c - partition* for X if it satisfies (Bezdek, 1981, p. 26):

1) $\mu_{ik} \in [0,1] \qquad i=1, \dots, c, \; k=1, \dots, n$

2) $\sum_{i=1}^{c} \mu_{ik} = 1 \qquad k=1, \dots, n$

3) $0 < \sum_{i=1}^{n} \mu_{ik} < n \qquad i=1, \dots, c.$

Elements can belong to two or more clusters to some extent, determined by the membership functions.

The location of a cluster is given by its centre

$$v_i = (v_{i1}, v_{i2}, \dots, v_{ip}) \in \Re^p \quad i = 1, \dots, n,$$

around which the elements are concentrated.

Then the objective functional is

$$z_m(U,v) = \sum_{i=1}^{c} \sum_{k=1}^{n} (\mu_{ik})^m \|x_k - v_i\|^2 \qquad (4.1)$$

Here v_i is the "centre of gravity" of cluster S_i, and x_k with a higher degree of membership have a higher influence on v_i.

$\| \, . \, \|$ is a norm defined on X and the distance between two elements is

$$d(x_k, x_l) = \|x_k - x_l\|.$$

The exponential weight m reduces the influence of the "noise" in the membership values, with relation to the clustering criterion. The larger is m, the more weight is assigned to elements with a higher degree of membership and the less to ones with a lower degree of membership. For $m=\infty$, U becomes

$$U=[1/c],$$

i.e. each element is assigned to each cluster with the same degree of membership. It is preferable, however, to have a less fuzzy U and usually $m=2$ is chosen.

Using differential calculus the necessary conditions for a local optimum for (4.1) are found to be (Bezdek,1981, p.67):

$$v_i = \frac{1}{\sum_{k=1}^{n}(\mu_{ik})^m}\sum_{k=1}^{n}(\mu_{ik})^m x_k \qquad m>1 \quad i=1, \ldots, c \tag{4.2}$$

$$\mu_{ik} = \frac{\left(\frac{1}{\|x_k - v_i\|^2}\right)^{1/m-1}}{\sum_{j=1}^{c}\left(\frac{1}{\|x_k - v_i\|^2}\right)^{1/m-1}}, \qquad i=1, \ldots, c \; k=1, \ldots, n \tag{4.3}$$

For all m>1 a fuzzy 'c-means' algorithm is designed that solves iteratively the necessary conditions (4.2) and (4.3) and converges to a local optimum (Bezdek, 1981, p.80).

4.2.4.2 The algorithm itself consists of the following steps:

Step1: choose c, $(2 \leq c \leq n)$;

m, $(m>1)$;

$G=[g_{ij}]_{pxp}$ a symmetric and positive definite *(pxp)* matrix, used to define a suitable norm in X

$$\|x\|_G = x'Gx,$$

where x' is the transpose of x. G indicates the relative numerical importance of each element and the correlations between the elements.

Examples of G are ***I**, the identity matrix, a diagonal matrix* ***[diag(g_{ii})]*** and the *covariance matrix* ***[cov(x)]***.

U is initialised as $U^{(0)}$ (the initial fuzzy partition, see section 3.2.2) and the steps counter l is set to 0, l=0.

Step2: Calculate the centres of the fuzzy clusters $\{v_i^{(l)}\}$ i=1,..., c using $U^{(l)}$ and equation (4.2).

Step3: Calculate $U^{(l+1)}$ using $v_i^{(l)}$ and equation (4.3) if $x_k \neq v_i^{(l)}$.

Else

$$\mu_{jk} = \begin{cases} 1 & for\ j = i \\ 0 & for\ j \neq i \end{cases}.$$

Step4: Calculate

$$\Delta = \| U^{(l+1)} - U^{(l)} \|_G.$$

If $\Delta > e$, then $l=l+1$ and go back to step 2. If $\Delta < e$ the algorithm stops; e is a small positive number, which measures the level of convergence.

4.2.4.3 The final problem for a clustering algorithm is to consider the validity of the clustering, which is "... the quality of the degree to which the final partition of a cluster algorithm approximates the real or hypothesised structure of a set of data"(Zimmerman, p.236).

This usually involves investigating only the appropriateness of c. The criteria, for both crisp and fuzzy cases, are discussed in Bezdek(1981).

4.2.4.4 The fuzzy c-means algorithm is widely used in many areas. There are many modifications, on the theoretical or on practical level, which are thought to do better in some special cases.

When applying this algorithm, there is no need to specify a probability distribution and the scale problems, considering characteristics which differ in magnitude, e.g. a simultaneous treatment of claim frequency and claim severity, can be treated through a suitably defined norm matrix.

Care is needed with the initial partition, because the algorithm comes to a local optimum and very different initial partitions may sometimes lead to non-consistent results.

There are no computational problems applying the algorithm even with a large number of elements.

4.2.5 An example which uses the above described fuzzy 'c-means' algorithm on a set of students from a postgraduate course in actuarial science is given in appendix 1.

4.3 Actuarial Applications.

4.3.1 The possible applications of pattern recognition in actuarial science are mainly in the field of risk classification. In this case, the aim is to group similar risks and to distinguish significantly different risks. Insurance risks and the criteria for determining the "good" risks are vague and ill-defined and therefore fuzzy clustering algorithms appear as natural candidates for handling this kind of data.

An advantage of the clustering in risk classification is that it does not need prior assumptions, but derives the clusters from the data. For example, as Ostaszewski (1993) states, in motor insurance in the US, where the rates are dependent on the driver's place of residence, charging the city inhabitants higher premiums could be interpreted by the public as discrimination against the inner city inhabitants and could be alleged to be racially motivated. Thus, it would be preferable to use methods that do not use any assumptions but rather discover any patterns in the data.

4.3.2 There are a few actuarial applications of fuzzy clustering. Two examples, using real insurance data, are given by Derrig and Ostaszewski (1995) and Yakoubov (1997) uses the above method for defining a grouping of a rating factor in motor insurance context.

4.3.2.1 Ostaszewski (1993) provides an example using artificial data. He considers a population of four people and each person being characterised by four features- weight, height, resting pulse and sex. The aim of the exercise is to form two clusters, one of them for standard lives and the second for substandard lives. The initial partition is based on sex and when the algorithm (fuzzy 'c-means') is performed[+] different clusters are formed, taking into account the remaining three features. The persons who are 'healthy' (lower weight/height ratio and lower resting pulse) form the first cluster and the "unhealthy" persons the second one.

4.3.2.2 The two examples given by Derrig and Ostaszewski (1995) look at the (a) the definition of motor rating territories and (b) the classification of motor claims with respect to possible fraudulent content. The motor rating data are drawn from the experience of the state of Massachusetts, USA. The State Commissioner for Insurance sets the premium rates for each rating territory in the state. In the study presented, 360 towns are identified for rating purposes and each town belongs to one of 24 rating territories. The towns are grouped by similar levels of expected losses, regardless of the geographic contiguity of the grouped towns, i.e the towns in a territory do not form a single connected area on the map.

The conventional approach uses an empirical Bayes procedure to determine a set of indices, one for each town, (DuMouchel, 1983), which are then ordered and cut-off points are chosen to determine the rating territories.

[+] Some minor errors exist, on p. 52, in the norm definition and on p.56, in the calculation of the entries for the first partition.

The 1993 indices are used, based on claim frequency data for five types of motor insurance claims over a four year period. Six rankings of the 360 towns are produced, one for each of the five types of claims and a combined index.

The assignment of the towns which are on the boundaries of the clusters to only one cluster is questionable. Furthermore, towns that reduce their relative claim cost, for example, by maintaining the roads, safety engineering codes or stricter law enforcement would expect a positive response from the rating system. The fuzzy algorithm detects fractional degrees of membership that may indicate that the town is in transition from one cluster to another and this can be used as an early warning of the change occurring. Compared to their current membership, towns may tend to become associated with nearby clusters as well as with their 'home' cluster. A geographical proximity factor can be included in the set of features. When it is combined with the expected claims, it results in fuzzier clusters, when applied to the whole state data. Neighbouring territories whose current territory boundaries become close could indicate a need for merging these clusters. Partial membership values for towns in a cluster could allow for the imperfect correlation between the different types of coverage.

4.3.2.3 The second application of the fuzzy 'c-means' algorithm, presented by Derrig and Ostaszewski (1995) deals with classification of the claims. It considers the growing problem of fraudulent claims, i.e. the claimants and/or the providers (medical services, garages) can manipulate the system for their own benefit.

Two types of fraud are distinguished:

-*criminal or hard fraud*, where there is a clear and wilful attempt to obtain money under false pretences and

- *build-up or soft fraud*, for example the ones that come from prolonged medical treatment or inflated car repair bills.

The information presented in support of a claim is vague and ill-defined, and it is difficult to find a benchmark, a model claim, because fraudulent claims are usually made to appear as normal and this makes the detection of fraud even more difficult. The studies of US general insurance data show that nearly 50% of the claims are suspicious, but only a very small part of these is detected and the appropriate actions taken (Weisberg and Derrig, 1992).

Ideally, one would like to construct a screening device that sorts the incoming claims into different "trays", according to their suspected fraud content. Attempts have been made to detect fraudulent claims, using regression models and neural networks. Because of the nature of the problem, fuzzy analysis has been considered as being a natural candidate for handling such data.

Derrig and Ostaszewski's investigation is based on a set of motor insurance claims data in Massachusetts, and the purpose is to determine how suspicion of fraud can be measured more precisely and to construct decision rules for determining the level of fraud content.

To investigate the ways of measuring the suspicion of possible fraud, for each claim a record on a zero to ten scale of suspected fraud is produced by a claims adjuster and four claims investigators. In order to "pool" more subjective opinions two more expert opinions are taken and a "vote" from "0" - none of the opinions classifies the claim as fraud, to "3"- three of the four opinions classify the claim as fraud (there was no one claim that collected four "votes"). Thus each claim is assigned with a three dimensional feature vector

(adjuster's suspicion, investigator's suspicion, result from the "voting").

The study shows that the clusters defined by the claims adjuster's level of suspicion approximately coincide with the fuzzy clustering algorithm applied to the 'better' information, represented by the feature vector, and thus the claims adjuster's expertise can be used to screen for possible fraud despite its subjective nature.

Having this result, the authors propose that a claims adjuster should evaluate each component of a particular claim and then an overall claim fraud assessment should be based on that evaluation.

4.3.2.4 Yakoubov (1997) uses clustering as a useful tool in the risk classification process for a motor insurer and describes three clustering algorithms, the minimum variance, an algorithm proposed by Loimaranta et al (1980) and the fuzzy c-means algorithm.

A grouping of the age factor based solely on claim data is presented using a fuzzy c-means clustering algorithm.

The idea behind the experiment is as follows.

A large number of underwriting factors in motor insurance are considered as factors, i.e. they take only a few values. The factors can have classes in the range of 2 as in the case of sex and more than 20 as for the car group. For example, driving experience is one of the underwriting factors in one of the states in the USA and it is grouped in three groups- 0-3, 4-6 and 7+ (Derrig, 1997). An obvious question is how the range of a rating factor can be divided into distinct and homogeneous groups, regarding the risk content, whether measured as a claim frequency, a claim severity or a pure premium.

The problem of age grouping is approached using a *fuzzy c-means algorithm*, based on past claim experience. The analysis is based on claim frequency data for two types of claims; material damage and bodily injury, and the relatively higher cost of the second claim type compared to the first is taken into account.

Each age is allocated to one or more clusters with a certain degree of membership. The so defined groups represent homogeneous risks and can be used further in the premium rating process.

4.4 Conclusions

Fuzzy pattern recognition and in particular fuzzy clustering methods are capable of giving realistic insights into problems with some degree of ambiguity. Thus Derrig and Ostaszewski conclude that fuzzy clustering, as applied to geographical risk classification, is able to identify fractional degrees of membership which may indicate towns or areas that are strongly associated with two or more clusters, Fuzzy clustering, as applied to fraud detection, provides a means of formalising the subjective claim assessment process, quantifying the components of the claim, instead of quantifying the claim itself. In defining groups over the range of a rating factor, fuzzy clustering provides a way to interpret fuzziness due to the inherent imprecision coming from the past experience data and to the possibility of an age being between two clusters. The age grouping under the fuzzy approach involves a substantial amount of judgement and would allow some other technical or marketing considerations to be included as well.

Appendix. THE ACTUARIAL EXAMINATIONS, fuzzy ' c-means' algorithm applied to determine the level of importance of mathematical background and degree level.

1. Introduction

To qualify as an actuary in the UK, the candidate is necessary to pass or gain exemptions from the professional examinations set by the Institute of Actuaries and Faculty of Actuaries. The first four subjects are technical and include the fundamentals of actuarial mathematics (subject A), economics, finance and accountancy (subject B), statistical methods in insurance (subject C) and more advanced actuarial mathematics (subject D).

The main requirement for prospective actuarial students is a first degree in a mathematical or related subject. As an illustration of the fuzzy techniques, the fuzzy ' c-means' algorithm is used to test and determine the level of importance of this entry requirement, i.e. do graduates in mathematics do better in the theoretical subjects A, B, C and D?

Since 1985 City University has offered a one year postgraduate diploma course in Actuarial Science, which covers the above professional subjects, effectively comprising eight half subjects (or their equivalents).

2. The experiments

2.1 Experiment 1

The data set in this experiment contains the exam results of 22 postgraduate students in Actuarial Science from City University, eleven of them having a degree in mathematics and/or statistics and remaining eleven having in economics, finance or some other non-mathematical degree.

Each student is assigned with a 8-dimensional vector $(x_{i1}, x_{i2}, \ldots, x_{i8})$ *and* x_{ij} is the i-th student's index result in subject j (subjects are ordered as A1, A2, B1, ... , for instance, subject A2 is the second element in the vector). The index result is defined to be the actual result of the student divided by the average for the subject and represents the student's relative performance in the subject. When a student is exempted from a subject, he/she is assigned an index result of 1 and the average for the subject includes only these students who have taken the subject. Table A1 presents the full data set.

The data set is divided into two equally-sized initial clusters, based on the first degree of the students. The students with non -mathematical or non- statistical degree are in initial cluster 1, and the remaining students are in initial cluster 2. Cluster i is determined by its characteristic function χ_i, defined as

$$\chi_i : S \to \{0\ ,\ 1\}$$

$$\text{and} \qquad \chi_i(s) = \begin{cases} 0 & \text{if student } s \notin \text{cluster } i \\ 1 & \text{if student } s \in \text{cluster } i \end{cases}$$

The elements belong to the initial clusters with certainty. The initial clusters are given in Table A2

The initial clusters are used as the initial fuzzy partition input for the algorithm (see section 4.2.4.1).

The norm matrix is the identity matrix I (we ignore the correlations between the subjects, although it is possible to calculate them from the data, and we allocate the same weights to each subject, although there may be 'easy' and more 'difficult' subjects).

A computer program in Visual Basic, run in Excel performs the iterative calculations of the fuzzy 'c-means' algorithm.

The results (the final fuzzy clusters) are given in Table A.3 and Fig. A.1 and Fig.A.2. Instead of only permitting full membership of a cluster, the fuzzy clustering allows partial membership. For instance, student 1, belongs to the second cluster with a degree of membership of 82% and to the first cluster with a degree of membership of 18%, instead of the 100% membership to the second cluster as we assumed at the outset.

It is surprising to see how well the final clusters correspond to the initial clusters (7 out of 22 cases deviate largely from the initial hypothesis), taking into account the fact that the degree type is only one of many and probably not the most important factor that may influence a student's performance in the examinations. Factors such as the degree level (first, second upper ....), student's own intellectual abilities and

and commitment, and the amount of work and time he is ready to sacrifice, are likely to be of equal importance as the type of the degree.

2.2 Experiment 2.

Let us adopt a slightly different approach towards the students with exemptions(from subjects B and C1). Instead of assigning an average mark, taking into account the fact that exemption level in actuarial examinations at City University is around 65%, we can recognise the exemption achievement by assigning a notional mark of 65%. Table A.4 represents the new index table, where 65% exemption mark is used in calculating the subject averages. The results from the fuzzy clustering are in Table A.5 and Figures A.3 and A.4.

The most dramatic changes are shown by three students-1, 4, and 18. The highest deviation (the membership value from section A.2.1 - modified membership value) is 15% and in two of the cases it is towards the initial clusters and in one of them it is in the opposite direction. If 50% is the required membership level for belonging to a cluster, the number of the misclassifications is reduced from 7 to 6, but the value for student 18 is only 60%, with value of 45% from experiment 1, therefore in both cases the student lies between the clusters.

The other changes are minor and not decisive for the final classification.

2.3 Experiment 3.

This experiment uses the data set from experiment 2 but investigates the relationship between the success in actuarial examinations and the grade of the first degree.

The set of students is divided into three groups: excellent students are in group 1 (with UK first or equivalent overseas grades), very good students are in group 2 (with UK upper second or equivalent overseas grades) and the remaining students are in group 3. Some of the data were incomplete and so the classifications were based on expected, rather than actual grade, obtained from letters of reference.

The three initial clusters and the final fuzzy clusters are given in Table A.6.

If we accept that an element belongs to the cluster in which it has the highest membership value, we have the following distribution of elements:

	group 1	*group 2*	*group 3*
initial grouping	5	11	6
fuzzy grouping	11	8	3

The final clusters differ quite significantly from our initial assumptions, only 8 out of 22 are in their initial groups. The reasons for this could be the following:

Firstly, the students come from different universities and countries with different systems and quality of education. Degrees of the same class may not be easily comparable.

Secondly , the fuzzy 'c-means' clustering algorithm divides the population into c 'equal' groups, i.e. the groups have equal weights. However, the grades, considered as groups are not equally weighted. For instance, it is not the case that achieving a first would require a score of between 67% and 100% on each dimension of work, intelligence, effort and so on, while achieving an upper second would require between 33% and 66% with for the lower grades requiring below 33%. The distribution of the qualities of students and the effort expected is unlikely to be 'uniformly' distributed in this simplistic matter.

Thirdly, the grade is unlikely to be the only important factor in determining future success in the actuarial examinations.

3. Conclusions

The experiment shows that the type of the first degree is an important feature which is associated with an actuarial student's performance at least in the first group of (technical) subjects, but it is not decisive. The partial memberships indicate the existence of other factors.

The grade of the first degree, while considered as very important, does not appear to be decisive for the particular data set examined. Looking at the result, one can deduce that students with not 'very good' first degrees are capable of doing well in the examinations. Indeed, four out of the six students with lower second or worse (initial group 3) are classified into the highest group (fuzzy group 1) and only one remained in group 3.

The fuzzy 'c-means' algorithm is useful in situations, where there is a need for grouping, based on a set of characteristics. The concept of partial membership is very convenient when the available information represents only a part of the true phenomenon, thus leaving the interpretation of the results to the judgement of the decision maker.

Table A.1 Set of index results for 22 students in 8 subjects, exempted subjects having index 1. Code 1 is for a mathematical or related degree and 0 is otherwise.

Students	Code	IndexA1	IndexA2	IndexB1	IndexB2	IndexC1	IndexC2	IndexD1	IndexD2
1	1	1.1384	1.0216	1.1446	1.0265	1.0000	1.0562	1.1736	0.9574
2	0	0.1708	0.0547	0.4077	0.4248	0.1492	0.3188	0.1401	0.3829
3	0	0.9534	0.7114	1.0000	0.5841	1.0000	0.7971	0.6306	0.7850
4	1	0.9392	0.8027	0.7369	0.9381	1.0000	0.9964	0.8232	1.1871
5	1	1.2950	1.3864	1.4111	1.2920	1.3610	1.3351	1.5589	1.3020
6	1	1.1242	1.0216	0.9878	0.9735	1.0000	1.3351	1.2962	1.1488
7	0	0.9819	1.0945	0.6899	0.8496	1.0000	1.4547	0.9984	1.1680
8	0	1.2238	1.3317	1.0662	0.8142	1.0000	1.3152	1.2787	0.9574
9	0	0.4981	0.6385	0.9721	0.6903	0.8576	0.6775	0.4379	0.7659
10	1	1.1527	1.1675	1.2544	1.4690	1.0000	1.4149	1.2436	1.3211
11	0	0.9677	0.7297	1.0000	1.1681	1.0000	0.7971	0.8057	0.5936
12	0	1.0530	0.9668	0.7213	0.5664	0.7458	0.5380	0.6481	0.4787
13	0	0.8538	0.7662	0.9251	1.1150	0.7085	0.6377	0.6131	0.5936
14	1	1.3946	1.6965	1.4582	1.5752	1.8271	1.8333	1.7341	1.7998
15	1	1.2523	1.3499	1.2544	1.2743	1.0000	1.1558	1.5414	1.1297
16	0	1.1953	1.1310	1.1446	1.4690	1.2305	1.1957	1.3838	1.3020
17	1	0.8680	0.7844	0.9878	0.6549	1.0000	0.8370	0.8232	1.0722
18	1	1.0530	1.3134	1.0000	1.0000	0.6712	0.9366	0.9108	0.7084
19	1	1.1384	1.0216	1.2700	1.1504	1.3610	1.0163	1.0510	1.2446
20	0	1.1100	1.1857	1.0000	1.0000	1.4356	1.2554	1.4188	1.3786
21	0	0.5265	0.3831	0.5801	0.6726	0.6525	0.2790	0.3153	0.3255
22	1	1.1100	1.4411	0.9878	1.2920	1.0000	0.8170	1.1736	1.3977
Average	0.5	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Table A.2 Initial clusters

Students	Code	cluster1	cluster2
1	1	0	1
2	0	1	0
3	0	1	0
4	1	0	1
5	1	0	1
6	1	0	1
7	0	1	0
8	0	1	0
9	0	1	0
10	1	0	1
11	0	1	0
12	0	1	0
13	0	1	0
14	1	0	1
15	1	0	1
16	0	1	0
17	1	0	1
18	1	0	1
19	1	0	1
20	0	1	0
21	0	1	0
22	1	0	1

Table A.3 The results from the fuzzy clustering

Students	Code	fyzzyclstr1	fyzzycltr2
1	1	0.181927	0.818073
2	0	0.79016	0.20984
3	0	0.907683	0.092317
4	1	0.593283	0.406717
5	1	0.069898	0.930102
6	1	0.09322	0.90678
7	0	0.296269	0.703731
8	0	0.141316	0.858684
9	0	0.941864	0.058136
10	1	0.066198	0.933802
11	0	0.785986	0.214014
12	0	0.896765	0.103235
13	0	0.923471	0.076529
14	1	0.201868	0.798132
15	1	0.058228	0.941772
16	0	0.043929	0.956071
17	1	0.762757	0.237243
18	1	0.54546	0.45454
19	1	0.121743	0.878257
20	0	0.066293	0.933707
21	0	0.879178	0.120822
22	1	0.153233	0.846767

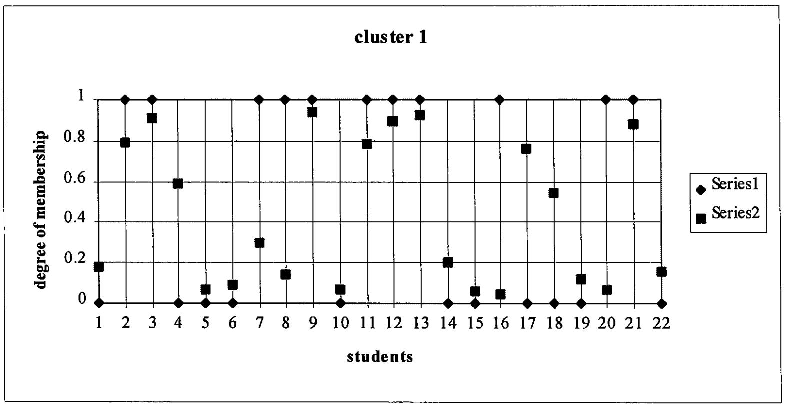


Fig.A.1 The initial membership to cluster 1is indicated by 1 and non-membership by 0.

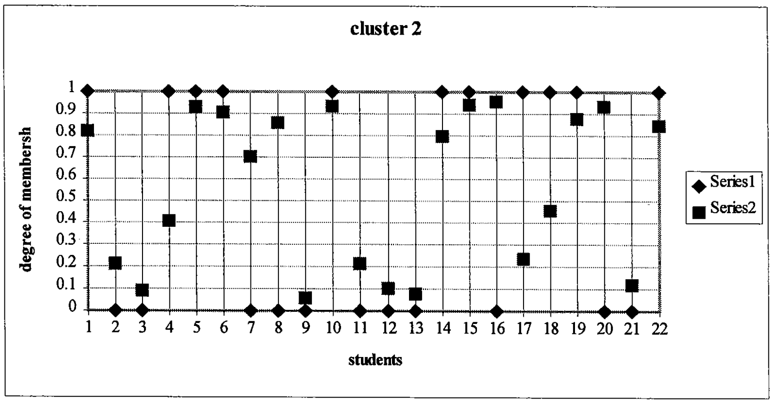


Fig.A.2 The initial membership to cluster 2 is indicated by 1 and non-membership by 0.

Table A.4

Students	Code	IndexA1	IndexA2	IndexB1	IndexB2	IndexC1	IndexC2	IndexD1	IndexD2
1	1	1.1384	1.0216	1.1406	1.012698	1.095785	1.0562	1.1736	0.9574
2	0	0.1708	0.0547	0.4063	0.419048	0.134866	0.3188	0.1401	0.3829
3	0	0.9534	0.7114	1.0156	0.57619	1.095785	0.7971	0.6306	0.7850
4	1	0.9392	0.8027	0.7344	0.925397	1.095785	0.9964	0.8232	1.1871
5	1	1.2950	1.3864	1.4063	1.274603	1.230651	1.3351	1.5589	1.3020
6	1	1.1242	1.0216	0.9844	0.960317	1.095785	1.3351	1.2962	1.1488
7	0	0.9819	1.0945	0.6875	0.838095	1.095785	1.4547	0.9984	1.1680
8	0	1.2238	1.3317	1.0625	0.803175	1.095785	1.3152	1.2787	0.9574
9	0	0.4981	0.6385	0.9688	0.680952	0.775479	0.6775	0.4379	0.7659
10	1	1.1527	1.1675	1.2500	1.449206	1.095785	1.4149	1.2436	1.3211
11	0	0.9677	0.7297	1.0156	1.152381	1.095785	0.7971	0.8057	0.5936
12	0	1.0530	0.9668	0.7188	0.55873	0.67433	0.5380	0.6481	0.4787
13	0	0.8538	0.7662	0.9219	1.10000	0.640613	0.6377	0.6131	0.5936
14	1	1.3946	1.6965	1.4531	1.553968	1.652107	1.8333	1.7341	1.7998
15	1	1.2523	1.3499	1.2500	1.257143	1.095785	1.1558	1.5414	1.1297
16	0	1.1953	1.1310	1.1406	1.449206	1.112644	1.1957	1.3838	1.3020
17	1	0.8680	0.7844	0.9844	0.646032	1.095785	0.8370	0.8232	1.0722
18	1	1.0530	1.3134	1.0156	1.134921	0.606897	0.9366	0.9108	0.7084
19	1	1.1384	1.0216	1.2656	1.134921	1.230651	1.0163	1.0510	1.2446
20	0	1.1100	1.1857	1.0156	1.134921	1.298084	1.2554	1.4188	1.3786
21	0	0.5265	0.3831	0.5781	0.663492	0.590038	0.2790	0.3153	0.3255
22	1	1.1100	1.4411	0.9844	1.274603	1.095785	0.8170	1.1736	1.3977
Average	0.5	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Table A.5

Students	Code	fc1	fc2
1	1	0.105886	0.894114
2	0	0.78388	0.21612
3	0	0.865532	0.134468
4	1	0.693152	0.306848
5	1	0.068254	0.931746
6	1	0.086624	0.913376
7	0	0.317617	0.682383
8	0	0.09561	0.90439
9	0	0.945811	0.054189
10	1	0.058783	0.941217
11	0	0.65603	0.34397
12	0	0.879832	0.120168
13	0	0.896087	0.103913
14	1	0.186818	0.813182
15	1	0.046023	0.953977
16	0	0.045206	0.954794
17	1	0.814863	0.185137
18	1	0.401941	0.598059
19	1	0.136265	0.863735
20	0	0.030778	0.969222
21	0	0.884123	0.115877

22	1	0.162422	0.837578

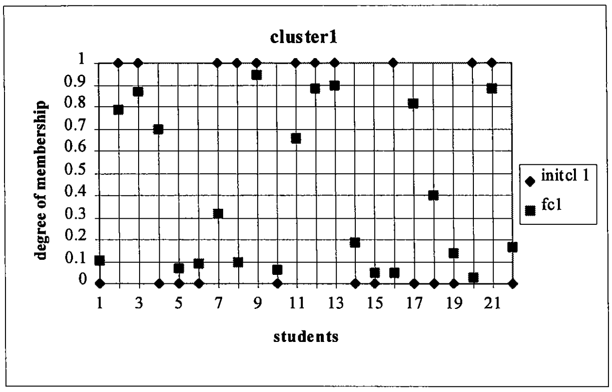


Fig. A.3

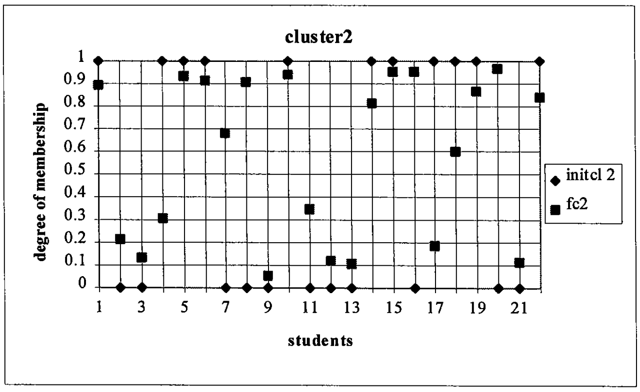
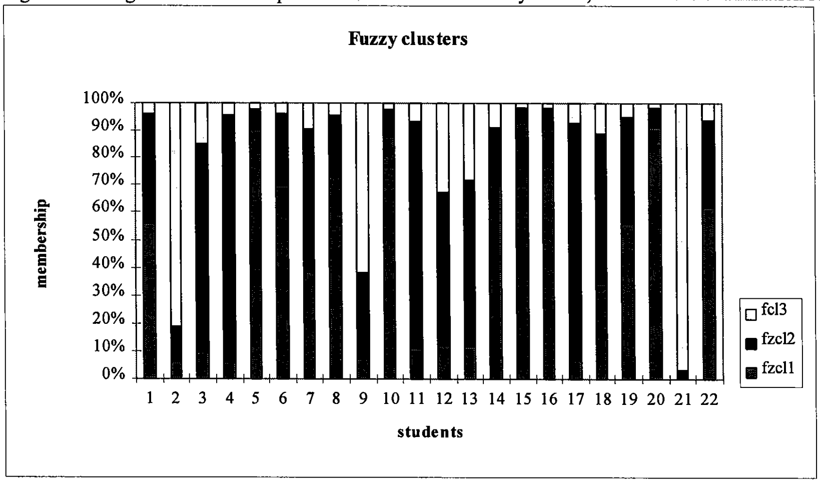


Fig. A.4

Table A.6

students	initcl1	initcl2	initcl3	fzcl1	fzcl2	fzcl3
1	0	1	0	0.558997	0.40084	0.040163
2	0	1	0	0.060764	0.128933	0.810303
3	0	0	1	0.093559	0.759044	0.147397
4	0	1	0	0.06269	0.89423	0.04308
5	0	1	0	0.898301	0.079427	0.022272
6	0	0	1	0.696659	0.262942	0.040399
7	0	1	0	0.381622	0.525055	0.093324
8	0	1	0	0.729786	0.226981	0.043233
9	0	0	1	0.070392	0.311643	0.617965
10	1	0	0	0.878949	0.09802	0.023031
11	0	1	0	0.108068	0.82437	0.067562
12	0	1	0	0.121403	0.553357	0.32524
13	0	0	1	0.115841	0.604605	0.279554
14	1	0	0	0.70061	0.21367	0.08572
15	0	0	1	0.927984	0.057682	0.014334
16	0	1	0	0.905517	0.077295	0.017188
17	1	0	0	0.067422	0.86181	0.070768
18	1	0	0	0.343746	0.5445	0.111754
19	0	1	0	0.559028	0.39361	0.047362
20	1	0	0	0.909424	0.075848	0.014728
21	0	1	0	0.00796	0.023653	0.968388
22	0	0	1	0.617643	0.319228	0.063129

Fig.A.5 The degree of membership of the elements to each fuzzy cluster, based on the examination results.



5. FUZZY EXPERT SYSTEMS AND FUZZY LOGIC CONTROL

5.1 Theory

5.1.1 An *expert system* is a computer program which holds knowledge of some subject and reasons with that knowledge with a view to solving a problem or giving advice. It can act as a human expert or play an assistant role in a decision making process. The origins of expert systems are in *Artificial Intelligence*, a branch of computer science which deals with the design and implementation of programs capable of problem solving, visual perception and understanding language. The general structure of an expert system is shown in Figure 5.1 (Zimmerman 1991, p.174).

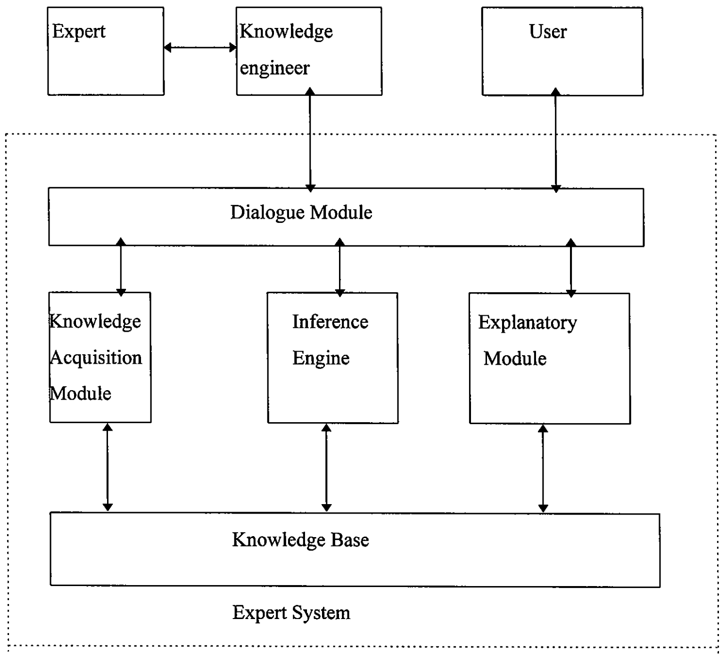


Fig.5.1 Structure of an expert system.

The knowledge about the event in investigation is entered in the ***knowledge base*** of the system via the ***knowledge acquisition module***. The knowledge base is made up of a *declarative part,* which describes the objects in the expert system and the relationships between these objects, and a *procedural part*, which contains information as to how these objects can be used to reach conclusions.

The expert knowledge is represented in the system in a format that best suits the domain of expertise of the system. The four most frequently used techniques are:

production rules, used mainly for procedural knowledge. The rules are in the IF ... THEN form:

IF a set of conditions is satisfied THEN a set of consequences can be produced. For example(Zimmerman, 1991, p.176)

IF the car won't start and the lights are dim

THEN the battery may be dead.

semantic nets, which are used mainly for declarative knowledge. The concepts are presented by a number of nodes, associated with one other by links. The links can represent various types of relationships between the concepts. Zimmerman(1991, p. 177) uses a semantic net to represent knowledge, related to motor vehicles. Part of the example is presented in Figure 5.2

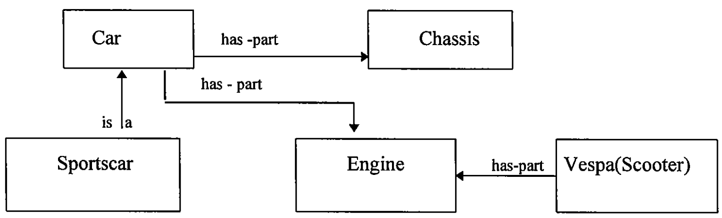


Fig.5.2 Semantic net

frames are introduced by Minsky (1975), as structures that collect together information about a concept and provide expectations and default knowledge about it. Typically, the frame is represented by a group of slots and their values, which themselves can be frames. For example part of the knowledge in the Fig.5.2 can be represented by a frame:

Frame:	car
Classes:	sportscar, small car, family car
Has parts:	chassis, engine;

and *predicate calculus* is a part of mathematical logic. In its arsenal, in addition to the conventional logic operators (and, or, ...) are included the *universal, or 'all'* operator ($\forall$) and *existential or 'some'* operator ($\exists$). For example, it allows us to express statements such as:

' $\forall$ *cars* have *chassis'*

' $\exists$ *vehicles* which have *engines*'.

The Inference Engine is a mechanism for forming inferences and drawing conclusions. The ways of deducing conclusions depend on the structure of the inference engine and the method used to represent the knowledge. In the case of production rules, two popular techniques are *forward chaining* and *backward chaining*. An extensive coverage of these and other inference techniques can be found in Waterman (1986).

5.1.2 Uncertainty is inherent in modelling expert systems. The following reasons for the use of fuzzy set theory are identified (Zimmerman, 1991, p.179) :

-The interfaces of the expert systems are with human beings, and therefore the use of the natural language as a means of communication involves the imprecision inherent to statements made by humans.

-The knowledge base of an expert systems contains human knowledge and it is usually the case that the rules and facts are neither totally certain, nor totally consistent. The storage of this kind of knowledge by using fuzzy sets seems very appropriate.

-The uncertainty in the knowledge base induces uncertainty in the conclusions and therefore there is a need to transfer the degree of uncertainty to the conclusions.

5.1.3 Fuzzy expert systems are structured as follows:

The *rules and facts* are presented using fuzzy sets. It enables the designer of the expert system to allow implicitly for uncertainty.

The *inference engine* is based on fuzzy logic and approximate reasoning methods. An example of fuzzy logic tool is *fuzzy modus ponens*, defined as the following expression:

Implication: If x is A then y is B

Premise: x is A'

Conclusion: y is B'

in which A, A', and B, B' are fuzzy sets defined in universes of discourse X and Y and x is a variable in X and y is a variable in Y. The fuzzy set A' measures the degree to which x 'satisfies' A and based on the degree of satisfaction, one defines the fuzzy set B'.

For example(Young, 1996), at the beginning of the year, an actuary might say,

'If the number of the insurance policies decreases by a *moderate* amount during the year, then I will decrease the rates *moderately.'*

This statement can be written as:

Implication: If the number of the insurance policies decreases by a *moderate* amount during the year, then I will decrease the rates *moderately.*

Premise: The number of the insurance policies decreased *somewhat moderately.*

Conclusion: I will decrease the rates *somewhat moderately.*

In this example x= *decrease in the number of policies*, y = *rate decrease*, A= *moderate decrease in the number of policies*, and B = *moderate rate decrease.*

5.1.4 Fuzzy expert systems have been developed and used in many areas of human activities. (Some of them are described in Zimmerman, 1991, chapter 10). There are systems in medicine used for diagnostic purposes, in earthquake engineering for assessing structural damages, in management and economics and in strategic planning.

5.1.5 *Fuzzy logic control (FLC)* systems are similar to *fuzzy expert systems*; both aim at modelling human decision making behaviour, but there are clear differences between the two systems. The origin of fuzzy logic control systems is in control engineering, rather than in artificial intelligence; its main characteristics are: the input mainly comes from a technological process; the decision making is rule-based and the output is a control statement.

The sequential steps in designing a FLC systems include (Zimmerman, 1991, p.204):

-Definition of the input variables and the possible control actions.

-Consideration of the way in which the observations are 'translated' into fuzzy sets and fuzzy control statements transformed into deterministic actions.

-Design of the rule-base, i.e. which rules under which conditions are to be applied.

- Design of the algorithm performing the necessary computations.

The applications of fuzzy logic controls vary from fuzzy control of a cement kiln, train operations control for the underground system in Sendai (Japan) to control of car by oral instructions (Zimmerman, 1991).

5.2 Actuarial applications.

5.2.1 The actuarial applications of fuzzy expert systems which have been suggested in the literature so far, are mainly in underwriting and risk classification, where the available knowledge and rules are vague and not very well defined. Very often, a rigid definition of some criterion, as 'a standard life has systolic blood pressure that does not exceed 130 mm of Hg' makes the distinction abrupt. It would be more sensible and useful to have a more flexible definition. Furthermore, the existence of two or more characteristics will have an influence on the output decision which is dependent on the extent and nature of each of the characteristics. For example, the life expectancy of a policyholder with a high blood pressure and a high ratio of actual weight to recommended weight is likely to depend on the level of these characteristics and their interaction.

5.2.2 A simple expert system that recognises a 'preferred policyholder' category is described by Lemaire (1990). In recent years, heavy competition between the American life assurers has resulted in a greater subdivision of policyholders and premium discounts being granted to policyholders who meet certain stringent health requirements-the 'preferred policyholder' category is a result of this refinement (Werth, 1995).

In Lemaire's system, a prospective policyholder x is characterised by four variables

$$x = x(t_1, t_2, t_3, t_4)$$

t_1 is the level of cholesterol in the blood, in mg/dl,

t_2 is the systolic blood pressure, in mm of Hg,

t_3 is the ratio of the effective weight to the recommended weight, as a function of height and build,

t_4 is the average consumption of cigarettes per day.

Each variable is assigned with a fuzzy set that describes the desirability of the criterion. For instance, if the normal systolic blood pressure is in the region of 130 mm and data from follow up studies demonstrate that a person with a blood pressure greater than 170 mm is approximately five times more likely to suffer from heart disease than an individual with a normal blood pressure. Therefore, Lemaire defines the fuzzy set of individuals with a normal blood pressure by the membership function:

$$f(x) = \begin{cases} 1 & x \leq 130 \\ 1-2\left(\dfrac{x-130}{40}\right)^2 & 130 < x \leq 150 \\ 2\left(\dfrac{170-x}{40}\right)^2 & 150 < x \leq 170 \\ 0 & x > 170 \end{cases}$$

Then, inference is based on appropriately defined intersections between the fuzzy sets. The definition of the intersection, depending on the nature of the characteristics, can allow for cumulative effects, interaction and compensation between the criteria. Having obtained the fuzzy set of 'preferred policyholder', then an α-cut can give the decision rule. For instance, if the intersection is the minimum intersection then an adopted 0.75 α-cut would give a decision at a point where each variable's membership function is at least 75%.

Furthermore, the approach of fuzzy set theory can accommodate differences in the level of importance of the criteria. If, for medical reasons, it is thought that the level of cholesterol is a better indicator of future heart problems compared to the level of blood pressure, then the fuzzy set that represents the cholesterol level can be concentrated, in order to reflect the importance of this criterion and the fuzzy set for blood pressure can be dilated(see section 2.2).

5.2.3 Another application of fuzzy expert systems to actuarial problems is given by Young(1993). The author describes two fuzzy expert systems in group health underwriting, where an employer offers single option or multiple-option health plans to its employees. The desirable characteristics of such a group, i.e. the rating factors include factors such as an appropriate *minimum number of employees*, a *minimum percentage of participation* in the plan, *stable morbidity*, ensured by a constant flow of young lives in the group, the *employer's involvement* in the plan, his *credit rating*, the *administrative structure, type of industry* (some are to be avoided), *ongoing claims should not be a large proportion* of the total claims, *good claims experience* and *low turnover rate* with respect to carriers.

Four categories of groups are considered: preferred risk, normal risk, substandard risk and unacceptable risk. Each criterion is associated wit a fuzzy set, that quantifies the level of risk. Young(1993) defines the boundaries of these four groups to be 1, 0.5, 0.25 respectively, i.e. a preferred group has a degree of membership 1 and an unacceptable group is below 0.25. For instance, the minimum participation factor is assigned with the following fuzzy set:

$$f(x)=\begin{cases}1, & 0.90\le x,\\ 5x-3.5, & 0.70\le x\le 0.90,\\ 0, & else\end{cases}$$

A preferred group is specified as having more than 90% participation, the normal group's participation is between 80% and 90% and a substandard group has participation of 75% to 80%.

The interaction between the criteria and the relative importance of one criterion to the others are modelled by an appropriate fuzzy intersections and fuzzy operations. After an analysis of interactions among the factors, the following fuzzy set is given as a solution to the underwriting problem

$$Q=[H(H(P,\ credit\ rating;0.5),\ turnover;0]^{1/3}\cap[participation\ x\ employers\ contribution]^{1/2}$$
$$\cap[H(participation,\ flow\ of\ lives;\ 0.5)]^{1/2}\cap[loss\ ratio],$$

where $H(A, B, p)$ is the Hamacher operator with parameter p and A and B are fuzzy sets (see 2.1.8). The parameter p determines the degree of interaction: $p=0$ brings mild interaction, and when p increases the degree of interaction increases, and for $p=1$ the Hamacher operator reduces to the algebraic product, which is the intersection with a maximum interaction effect.

P is a linear combination of the administrative function factor, type of industry factor, and on-going claims factor.

Participation in the plan and employer's contribution have maximum interaction and they intersect through the algebraic product. The square root makes this term commensurate with the others.

The loss ratio does not interact with any of the other variables; it is a single term.

The cut - off points for Q determine the underwriting result for a particular group. For example if for a group, Q lies between 1 and 0.75, then the group represents a normal risk.

A sensitivity analysis of the system, i.e. how changes in the input variables affect the output, can be carried out in order to verify that the chosen functions characterise the qualities appropriately and that the combination of those functions accurately reflects the given underwriting process.

A similar, slightly more complex model is proposed by Young (1993) for describing the underwriting of multiple-option plans.

5.2.4 A further example is provided by Hellman(1995) who describes a fuzzy expert system which evaluates Finland's 461 municipalities with respect to the interest they represent to the insurance industry. The problem is not to find the rich and big provinces or small ones, but those of average size that are well managed and whose insurance cover is not adequate. The recognition of these provinces could lead to wider marketing efforts by the insurer and, it is hoped a gain of a new business.

The selection factors that are of interest to the insurer are: the population of the province, the non-life premium income, claims ratio, whether the province is already a client, the structure of the province's finances, the level of capital expenditure, the growth potential and the province's current financial strength.

After consulting an expert about the impact of these factors, Hellman describes them using fuzzy sets with a membership function that is a smoothed variant of the expert opinion. For example, the claims ratio factor(including the investment income) has the following membership function, determined by the expert:

claims ratio	*degree of membership*
less or equal to 15%	1
25% claims ratio	0.97
35%	0.93
50%	0.85
75%	0.65
100%	0.05
greater or equal to 122%	0

and the smooth function that is used to represent the membership function for the fuzzy set is

$$f_c(x) = \min\left\{1, \max\left\{0, \frac{43(1-\arctan\left(\frac{100x-85}{18}\right))}{100}\right\}\right\}$$

Having fuzzy sets for each factor, the next step is to combine the fuzzy sets in a way that most accurately represents the problem. Hellman divides the set of factors, described above, into three groups- economic, insurance and adjusting groups of factors. Then, each group is considered on its own merits. For example, the following structure for the group of economic factors (population, financial structure and level of capital expenditure) is proposed:

$$f_{econ} = 0.5 f_{population} + 0.5(f_{financial\ structure} \otimes f_{capital\ expenditure}),$$

where f stand for the respective membership function and $\otimes$ is the normal multiplication operation (or in terms of fuzzy sets it corresponds to the algebraic product, see 2.1.8). The multiplication in the above formula means that if both factors have low values, their combined effect is even more severe, and the equally weighted summation means that the effects of the population factor and the combination of the two financial factors is thought to be of similar importance.

5.3 Conclusions

The advantages of the fuzzy approach to expert systems over the traditional models are in (a) the simplicity of the models, which makes them easy to understanding; in (b) the ability of the models to accommodate vague, imprecise and subjective knowledge and therefore provide useful techniques in areas such as customer evaluation and marketing; and in (c) the flexibility of their structure, expressed by the ease with which the shape of the membership function can be changed or new factors thought to play a significant role in understanding a real-world problem can be incorporated.

Fuzzy modelling allows the simultaneous treatment of the problem being investigated and the model for solving this problem and therefore a much more realistic and closer relationship between the problem itself and the model can be established and utilised.

It can be argued that fuzzy expert systems are subjective and the output depends to a great extent on the input to the system. However, the expert opinion, on which the description of the system is based, could be thought of as the best available knowledge about the problem. Similarly, the opinions of several experts can be pooled in order to come to more objective decisions.

The construction of a 'good' expert systems requires skill and a thorough knowledge of the phenomenon, which is modelled and its interactions with the other processes.

Fuzzy expert systems can become an important part of so called 'hybrid' systems, which use a combination of approaches to model a phenomenon as accurately as possible. A more detailed discussion of such systems is presented in section 7.

We believe that fuzzy expert systems, with their superiority in dealing with the presence of information of poor quality and the facility for automating the decision process, are one of the approaches with considerable future potential in the areas of risk classification and underwriting. A promising example of this is the ready-to-use fuzzy expert system presented by Horgby et al (1997) which quantifies the risk of an applicant with diabetes mellitus for a life policy.

6. DECISION MAKING IN FUZZY ENVIRONMENTS

6.1 Theory

6.1.1 In statistical decision theory, a decision is characterised by a set of *decision alternatives* (the decision space), described by enumeration or by a set of constraints; a set of *states of nature* (the state space); *a relation* assigning to each pair of decision and state a result and *a utility function* that orders the results according to their desirability.

Fuzzy decision theory considers the situation where the utility (objective) function and constraints are fuzzy and both objective and constraints are modelled by fuzzy sets. Since by analogy to the non-fuzzy case, we want to satisfy simultaneously the objective and the constraints, therefore a decision can be viewed as the intersection of the fuzzy constraints and the fuzzy objective(s). This then provides a fully symmetric relationship between the constraints and the objectives.

6.1.2 Let us consider an example. The board of directors wants to determine the optimal dividend to be paid to the shareholders of the company. It ought to be attractive and for reasons of public and employee relations it should be modest. An 'attractive' dividend can be defined as the fuzzy set with membership function f_A(Zimmerman, 1991, p. 244)

$$f_A(x) = \begin{cases} 1, & x \geq 5.8 \\ \dfrac{1}{2100}(-29x^3 - 366x^2 - 877x + 540) & 1 < x < 5.8 \\ 0 & x < 1 \end{cases}$$

The so defined function is *0* for $x<1$, i.e. dividend less than 1% is not attractive at all, *1 for* $x>5.8$, i.e. this is the domain of the attractive dividends and between 1 and 5.8 f is an increasing cubic function.

The fuzzy set 'modest dividend' can be defined by f_M

$$f_M(x) = \begin{cases} 1, & x \leq 1.2 \\ \dfrac{1}{2100}(-29x^3 - 243x^2 + 16x + 2388) & 1.2 < x < 6 \\ 0 & x \geq 6 \end{cases}$$

The definition implies that a dividend is 'modest' for certain if is less than *1.2%,* it is not 'modest' at all if it is greater than *6%* and between *1.2%* and *6%* is modest to some 'decreasing' degree.

Then the fuzzy decision would be characterised by its membership function $f_D=min(f_A, f_M)$ and if the decision maker wants a 'crisp' decision then the obvious choice would be the element with highest membership value.

$$x_{max}=arg(max\ min(f_A(x),\ f_M(x)),$$

where *arg* stands for the argument of the function, i.e. arg(f(x))=x. The so defined decision is called a 'maximising decision' and is similar to the conventional maximin criterion.

6.1.3 Formally, a decision in a fuzzy environment in the sense of Bellman and Zadeh (Bellman and Zadeh, 1970) is defined as:

if $X = \{x_1, x_2, ..., x_n\}$ is the set of alternatives

$G_1, G_2, ..., G_p$ are the fuzzy goals, represented by fuzzy sets

$C_1, C_2, ..., C_q$ are the fuzzy constraints,

the decision D is defined to be the fuzzy set

$D = G_1 * G_2 * ..., *G_p *C_1 *C_2 * ... *C_q$ where * are appropriately defined, probably context dependent, connective operators.

The set of the decisions can be found as

$$K = \{ x \in X : \mu_D(x) = max \} .$$

Each alternative that belongs to the set is an optimal decision. This is one of the approaches in choosing optimal decisions. The problem is context and goal dependent. A discussion of some alternative approaches is given by McCauley-Bell and Badiru (1996).

6.1.4 Special kinds of fuzzy decision models are the fuzzy linear programming models (FLP). The classical LP problem can be stated as :

$$
\begin{array}{ll}
\max & c^t x \\
\textit{such that} & Ax \le b \\
& x \ge 0 \\
\multicolumn{2}{l}{\textit{with } c, x \in R^n, b \in R^m, A \in R^{mxn}}
\end{array}
\qquad (6.1)
$$

Possible fuzzifications of the classical LP would be in terms of:

-a fuzzy objective function, e.g. one that describes a statement such as 'improve the present cost situation considerably';

-vague constraints, the relation *(<)* or the parameters *(A, b, c)* are fuzzy , because of their nature or because the perception of them is fuzzy.

6.1.5 The simplest FLP model is the *symmetric FLP*, as described by Zimmerman (1991, p.250). It is assumed that the decision maker can establish an aspiration level, z, for the value of the objective and each of the constraints is modelled as a fuzzy set. The FLP model becomes:

$$
\begin{array}{ll}
& c^t x \succ z \\
\textit{such that} & Ax \prec b \\
& x \ge 0 \\
\multicolumn{2}{l}{\textit{with } c, x \in R^n, b \in R^m, A \in R^{mxn}, \prec \textit{ is the fuzzified version of } \le.}
\end{array}
\qquad (6.2)
$$

A fuzzy inequality relation can be described as 'essentially smaller/greater than or equal' (see section 2.4). The first two inequalities in (6.2) are fully symmetric (the objective and the constraints) and we can combine them:

$$
\begin{array}{l}
Bx \prec d \\
x \ge 0
\end{array}
\qquad (6.3)
$$

where *B* and *d* are respectively an appropriately defined matrix and vector. Each of the *(m+1)* rows are represented by a fuzzy set and if the membership functions are f_i, then the decision *D* is given by the fuzzy set with

$$
f_D(x) = \min_i \{ f_i(x) \} \qquad (6.4)
$$

$f_i(x)$ can be interpreted as the degree to which x satisfies the fuzzy inequality $B_i x \prec d_i$ (where B_i is the i-th row of B).

6.2 Actuarial applications.

6.2.1 Lemaire(1990) gives an example that uses the fuzzy decision procedure to identify the optimal excess of loss retention level of a reinsurer that offers 10 different deductibles.

The goals and constraints for an optimal decision are derived from ones used in reinsurance practice and their nature is imprecise and vague.

Four variables are evaluated for each level of retention: the probability of ruin, the coefficient of variation of the retained portfolio and the respective ratios of the reinsurance premium and the deductible to the original premium income.

For instance, it is plausible for the *reinsurance premium* not to exceed 2.5% of the premium income (of the line of business) by *much*. This statement can be modelled by a fuzzy set and Lemaire suggests the following membership function, represented in Figure 6.1:

$$f(x,t)=\begin{cases}1 & t\le 2.5\\ 1-2(\frac{t-2.5}{0.6})^2 & 2.5<t\le 2.8\\ 2(\frac{3.1-t}{0.6})^2 & 2.8<t\le 3.1\\ 0 & 3.1<t\end{cases}$$

where x is an alternative (see section 6.1.3).

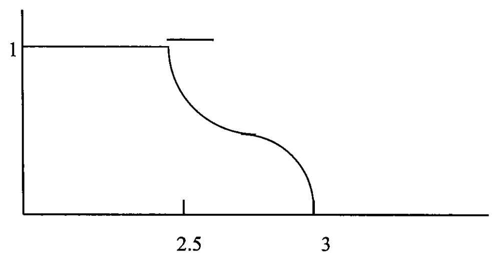


Fig. 6.1 The membership function of reinsurance premium/cedents' premium income

The membership function of the fuzzy decision is then easily calculated from the functions for *the goals*- low probability of ruin, not exceeding 0.00002 and low coefficient of variation, not exceeding 3.1 and

the constraints - reinsurance premium less then 2.5% and the deductible close to a level of 1% of the line's premium income.

The alternative with the highest degree is chosen as the optimal decision of the problem.

It is important to mention that the goals and constraints may be conflicting and it may not be possible for all of them to be fully satisfied. It could be decided that one of them is much more important than the others and that higher weight should be given to the membership function using the concentration operation. Also different definitions for intersection could be used when small violations of constraints are acceptable or just some improvements in the objective function (rather than the maximum) are required. Fuzzy decision making allows these features to be incorporated explicitly in the model.

6.2.2 Cummins and Derrig (1993) look at the accuracy of forecasting pure premium rates. Claim cost forecasting models are developed and a fuzzy approach is used to choose the best method. To the standard actuarial methods are added econometric and more sophisticated time trend models. Each forecasting method is characterised by its estimation period, estimation technique, frequency and claim size models. The objective is to apply the best of the existing methods to forecast pure premiums and then to use a fuzzy decision procedure to solve the model selection problem.

The measure of accuracy of a model is the *total predicted change error* (TPCE), defined as:

$TPCE = (1+d)/(1+d_a) - 1$

$1+d$ is the predicted trend factor and

$1+d_a$ is the actual trend factor.

72 forecasting methods are used to estimate the trend factor. The primary measure of the forecast performance is the *average absolute TPCE* and the *average TPCE* is used as an alternative accuracy measure mainly as an indicator of the degree of bias, i.e. the tendency of methods consistently to over-predict or under-predict the trend.

The experiment showed wide variation in the estimates, even among the "best" methods. The conventional choice of "best" forecasting method would be the one with the optimal value under certain specified statistical decision criteria. Instead a fuzzy decision procedure is used.

X, the *set of alternatives* is the set of 72 forecasting methods. The *objective* of the decision making is defined as a reasonable closeness to the average of all methods, i.e. the chosen best method should not be extreme, compared to the 'average' method.

The *constraints* imposed on the choice come from *historical accuracy*, whether the method was good for past predictions measured by the average absolute TPCEs and *unbiasedness*, measured by the average TPCEs.

Fuzzy decision making assigns to each method a degree to which the method is good for the decision maker. If the method with the lowest absolute average TPCE is labelled the *best method*, then the fuzzy set that describes the first constraint, the historical accuracy has a membership function *U(x),* defined as

$$U_1(x) = \frac{TPCE(Best)}{TPCE(x)},$$

the second constraint is made operational by defining the membership function

$$U_2(x) = e^{-|Average(TPCE(x))|}$$

and finally, the objective of a 'moderate, not too extreme' method is represented by the fuzzy set with the membership function

$$U_g(x) = \begin{cases} 1 - \frac{|d_x - \bar{d}|}{2\sigma}, & if \frac{|d_x - \bar{d}|}{2\sigma} \leq 1 \\ 0, & otherwise \end{cases}$$

where σ is the standard deviation of the trend factors, d is the mean value of the trend factors and d_x is the trend factor produced by the method. Under an appropriately chosen intersection operator, the degree of membership of each method is calculated. The best forecast is determined either

as the one which results from the method with the highest degree of membership or

as the weighted average of either the entire set of trends or of an α-cut, where the weights are the degrees of membership of the methods.

This second approach, as Cummins and Derrig (1993) mention, allows information from the closely ranked forecasting methods to be taken into account and it is not very sensitive to the choice of an intersection operator, which is not the case with the maximum membership approach. The experiment shows that the weighted average trend factor (from the second approach) is not very different from the trend factor with the highest degree of membership.

6.2.3 In their work Guo and Huang (1996) look at a modified version of the mean-variance asset allocation method, where skewness is included in the model.

The formulation is as follows. If there are N asset classes with allocation weights x_i and the rate of return for asset i is R_i , a random variable, with first and second moments, respectively μ_i and σ_i and R_p is the rate of return for the whole portfolio, then the following programming problem can be set up:

$$\begin{array}{ll} Min & \sum_{i=1}^{N} x_i^2 \sigma_i^2 + \sum_{i=1}^{N}\sum_{j=1}^{N} x_i x_j \sigma_i \sigma_j \\ Max & E[(R_p - \mu_p)^3] / \sigma_p^3 \\ Subject\ to & \sum_{i=1}^{N} x_i \mu_i = \mu \\ & \sum_{i=1}^{N} x_i = 1 \end{array} \qquad (6.5)$$

i.e. the conventional minimising of the variance of the rate of return of the whole portfolio, while the expected rate of return for the portfolio is fixed at some desired level.

The rationale behind the second objective, maximising the skewness of the rate of return is that we prefer small downside risk, i.e. large positive skewness indicates 'fatter' right tail, i.e minimising the risk of lower return (the variance effect) and maximising the 'risk' of higher return (the skewness effect) for a given expected rate of return. As stated by Guo and Huang, in a continuous time model, moments higher than two are irrelevant to the asset allocation decisions (this follows from Ito's differentiation rule for assets whose prices follow a diffusion process), but in the discrete case the higher moments are relevant to the manner in which assets are allocated.

If R_i are assumed to be triangular fuzzy numbers, i.e. $R_i = TFN(r_i^p, r_i^m, r_i^o)$, where r_i^p is the most pessimistic value for R_i, r_i^o is the most optimistic value and r_i^m is the most likely value, then (6.5) is approximated by (Guo and Huang, 1996)

$$\begin{aligned} &\min \quad \sum_{i=1}^{N} (r_i^m - r_i^p) x_i \\ &\max \quad \sum_{i=1}^{N} (r_i^o - r_i^m) x_i \\ &\textit{Subject to} \qquad\qquad\qquad\qquad (6.6) \\ &\sum_{i=1}^{N} r_i^p x_i = \beta \\ &\sum_{i=1}^{N} x_i = 1 \end{aligned}$$

Problem (6.6) is a multiple-objective linear programming problem which can be solved by using fuzzy programming methods.

The incorporation of the skewness in the model leads to a non-linear programming with the associated computational complications. Fuzzy programming provides an alternative way of representing and solving problems which include uncertain rates of return. It reduces the complexity of the multiple-objective programming problem (6.5), preserves the linear structure of the classical mean/variance model and provides an approximate solution.

6.3 Conclusions

Fuzzy set theory provides a systematic approach to decision making problems. Vague and conflicting requirements can be accommodated in order to produce results that satisfy these constraints to some extent.

The choice of membership function and intersection operator is subjective and allows for the decision maker's judgement to be incorporated. The fuzzy decision making procedure provides an effective method for mixing statistical results with the decision maker's judgement.

Fuzzy decisions are 'consensus' decisions, because they allow simultaneously for satisfying multiple objectives and constraints, and this may explain why they have flourished in Japan.

The fuzzy linear programming approach of Guo and Huang is a type of alternative to the mean/semi-variance model developed by Markowitz, as a modification of the standard modern portfolio theory. It emphasises the fact that deviation above the mean is not a risk to which the investor should be averse. It would be interesting to see how the results from the above approach compare to ones obtained using mean/semi-variance model, for example, and this is left to future research.

7. FUZZY TECHNIQUES AS PART OF 'HYBRID' MODELS

7.1 Introduction

7.1.1 A prevailing opinion among statisticians and actuaries is that fuzzy set theory aims to replace probability theory in modelling financial and risk management processes.

7.1.2 There are concepts which are inherently non random, for example the degree of disability can vary from 0% to 100% and it can not be described by a random variable (Hellman, 1995); or some other variable can be random and fuzzy at the same time, such as the claim amounts, arising from a general insurance business (Cummins and Derrig, 1997). In practice, future claims cost is predicted, based on the past experience data, which can have different levels of reliability. In addition, one may want to allow implicitly for inflation or some social changes (for example higher court awards in personal injury cases). Therefore, the claim amount can be thought of as a hybrid number, consisting of a random part and a fuzzy part, where the random part represents the realisation from the random process and the fuzzy part accommodates the influence of some other factors such as data reliability, inflation and some other factors that are not easily quantifiable.

7.1.3 Cummins and Derrig(1997) state:

> 'FST actually does not compete with mathematical probability as a means of evaluating random phenomena, but rather is complimentary to probability theory in dealing with real world problems where the available information is subjective, incomplete, or unreliable.'

7.1.4 One of the new and promising ways of using the fuzzy techniques is in combination with other deterministic and statistical methods. Fuzzy set theory provides a rigorous, mathematical way of utilising vague, subjective and qualitative concepts which are difficult to be incorporated into other methods.

7.2 Examples of hybrid systems

7.2.1 An actuarial example of a hybrid system which uses fuzzy techniques has been developed by the Dutch Insurance Supervisory Bureau, ISB (Kramer, 1997,p. 191, footnote) called "Early". It is an early warning system for detecting possible difficulties of pension funds and it is a combination of a logit model and a fuzzy set model, based on analysis of financial ratios of pension funds.

7.2.2 A similar hybrid system is developed by Kramer (1997) and it is an early warning system for evaluation of the risk exposure of non life insurance companies. Kramer's system, NEWS, combines a traditional statistical method, an ordered logit model with a neural network and an expert system. Although it does not include fuzzy techniques, we believe that if they had been included in the system, it would have given interesting results. In the following paragraphs, we will look at the original model and will suggest some modifications.

The classification of the insurance companies is based on the actual assessment made by the supervisory body, which distinguishes between high risk, medium risk and low risk companies. The data set upon which the system is based consists of seventy firm-specific variables, which cover the solvency, profitability, investments, reinsurance, types of risks insured, technical provisions, premium growth, size, and dependence on company groups: these factors are available for all insurance companies in Holland, except the relatively new companies and those companies in a run-off situation.

7.2.2.1 An ordered logit model is used to rank a company as a low, a medium or a high risk. This is a place where fuzzy techniques can be used. A fuzzy clustering algorithm can be used to classify a company to one of the above-named categories. The advantages of such approach are the possibility of partial memberships of the categories, the possibility of weighting the variables according to their influence and a flexible interpretation of the results.

7.2.2.2 The second stage of NEWS uses a 'specially' trained neural network (which can be considered as a special type of nonparametric regression estimator (Geman et al. 1992)) to obtain a second classification of the companies.

The overall result (where a misclassification by one method is compensated by a correct classification by the other method) is a combination of the results of the two methods. Although the results, concerning the high and low risk are very good (over 96% are correctly classified), both methods fail to recognise the medium risk group. Again it is our belief that an appropriate fuzzy clustering method could solve this particular problem.

7.2.2.3 The combined results from the logic and neural network methods are used as inputs to the final stage of the system, the expert system. Although the analysis of the financial ratios is important in the final assessment of a company, the expertise of the supervisor also needs to be included. The output is in the form of priorities for investigation for each company. Because the supervisor's expertise is not in quantitative form and very often is not strictly defined, we believe that the best representation of this expertise would be given by fuzzy sets. Therefore, we suggest that it would be worth exploring the idea of using a rule-based fuzzy expert system, where the rules are based on fuzzy approximate reasoning methods.

7.2.3 Two interesting hybrid models, which combine an Analytic Hierarchy Process (AHP, see the Appendix to this section) methodology with Fuzzy Set Theory (FST) are presented by Chen&He (1997) and McCauley-Bell&Badiru (1996). Neither application is actuarial in content-both deal with disability related issues which nevertheless have some relevance to potential actuarial applications.

7.2.3.1 Chen&He (1997) present an assessment model for measuring and interpreting a disabled person's (visually impaired in their study) capabilities to comply with the requirements of a work environment. AHP is used to obtain a hierarchical framework for measuring an Overall Disability Index (ODI). The index consists of factors, such as achievements, aptitudes, interests and opinions and other factors. Each of them is treated as a subindex and is measured by one or more evaluation systems. For instance, the *'interests and opinions'* subindex in the model is measured by three tests- *WRIOT* (Wide Range Interest Opinion Test), *AAMD* (American Association of Mental Deficiency) and *COPS* (Chen&He, 1997, p.4). The structure which follows from WRIOT is given in Figure 7.1

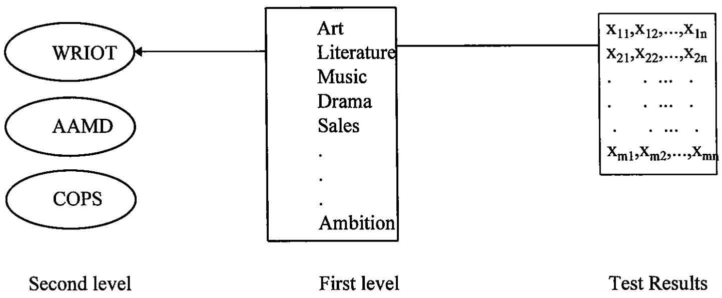


Figure 7.1 WRIOT evaluation system.

Then the first level of the AHP consists of elemental test results, which give input data to the evaluation system, which is one of the evaluation systems in the second level. A particular subindex is measured by the evaluation systems in its second level. Finally the Overall Disability Index is a linear combination of the subindices.

Developing such a model is a Multiple-Attribute Decision-Making(MADM) problem, where the weights for each attribute are to be defined. The eigenvector method and entropy method are used to derive the weights at the first level of the AHP. A single measure from the evaluation systems is not considered appropriate, because of the imprecision and the uncertainty of the results, due to the expert's judgements, evaluator's measuring skills and person's feelings during the test (Chen&He, 1997, p.8).

The solution which is offered is to treat the results as fuzzy sets. Then a subindex is defined as a weighted intersection of the fuzzy sets for the evaluation systems. And finally the ODI is a weighted sum of the subindices. Therefore, the systems defines the set of disabled people as a fuzzy set and each person belongs to the set (i.e. is disabled) to a certain extent, which is determined by the membership function.

7.2.3.2 McCauley-Bell&Badiru (1996) develop a similar model, based on AHP and FST. The model looks at the quantifying the significance of the risk factors for cumulative trauma disorders (CTD) of the forearm and hand and their impact on the likelihood of such injuries occurring. The research focuses on the use of linguistic variables as determinants of the risk levels, and quantifies these variables using FST.

AHP is used to obtain the weights for the factors in each of three categories, which were identified in a preliminary analysis and which are considered to constitute the overall risk. The categories of risk factors are: task-related, personal and organisational. For example, the following are the AHP results from level 1(McCauley-Bell& Badiru, 1996, p.134):

Task related Risk		Personal Risk		Organisational Risk	
Factor	Relative Weight	Factor	Relative Weight	Factor	Relative Weight
Joint posture	0.299	Previous CTD	0.383	Equipment	0.346
Repetition	0.189	Hobbies and habits	0.223	Production rate	0.249
Hand Tool Use	0.180	Diabetes	0.170	Ergonomics	0.183
Force	0.125	Thyroid prob.	0.097	Peer influence	0.0645
Task duration	0.124	Age	0.039	Training	0.059
Vibration	0.083	Arthritis	0.088	CTD level	0.053
				Awareness	0.045

Fig. 7.2 AHP, level 1.

The weights for each of the three risk groups add to 1 and they determine the relative impact of each one of the factors in the three categories to the output. The second level consists of defining the overall risk, which is determined by the three categories. The relative importance of the categories is given in Figure 7.3:

Category	Relative weight
Task related Risks	0.637
Personal Risks	0.258
Organisational Risks	0.105

Fig. 7.3 AHP, level 2.

The linguistic variables and the weights obtained from the AHP are used to evaluate the current condition of a given category using fuzzy inferencing (a rule-based expert system) and the output is a linguistic variable, such as 'very high risk', 'high risk', 'average risk', 'some risk' or 'little or no risk'. Again a weighted sum of the categories is considered most appropriate.

For example within task-related risks, if the factors *force, repetition* and *joint posture* have values 'very high risk' or 'high risk', then the fuzzy inferencing rules that the output for task- related category is 'very high'. The overall risk is a function of the linguistic output from the three categories. The complete system is realised in the model CTXPERT, developed by the authors.

The defuzzification, the process of translating the fuzzy outputs into a crisp value, follows Hayashi's 'fuzzy quantification theory' (Terano et al, 1987) and more precisely Theory I, which attempts to find a linear relationship between qualitative and numerical variables, assigning a numerical value to the linguistic variables.

The system is tested on a data set and *actual* vs *predicted* values are considered. Sensitivity, the ability to identify the CTD cases and specificity, the ability to identify only CTD cases analyses (types I and II errors) indicate good adherence to the actual experience. This is confirmed by a χ^2 test.

As noted by McCauley-Bell& Badiru(1997), FST provides a useful approach to the assessment of injuries or disabilities, which can develop over a period (possibly prolonged). FST can be used to measure the degree to which an injury or disability occurs and not just whether it occurs. Thus, 'the representation of this grey area is critical to recognise whether an injury is developing to make efforts to prohibit the continual exposure to the risk factors'.

7.2.4 The two hybrid systems (described in section 7.2.3) which combine knowledge, AHP and fuzzy inference show very good results. Furthermore, the models considered in this section are an indication that FST is a reasonable and effective approach in assessing risks, that are characterised by great complexity and variability. Although these are not actuarial models, one can easily see applications of such or similar models in situations with a great variability of individual factors in an insurance or risk management context, for example underwriting and risk classification, and the pricing of general insurance risks. Potentially, these methodologies could lead to the replacement of the underwriting manuals that are currently in widespread use.

7.3 Conclusions

7.3.1 Hybrid models provide an interesting approach to the complexity of real world applications. Various theories and methodologies may represent better different sides of an investigated phenomenon in a more effective way. A carefully chosen combination of these can be used as a complement to those aspects of a traditional method which are poorly modelled or as a correction or a double-checking device.

7.3.2 We believe that the development of hybrid models is a promising area for further research.

Acknowledgement

The authors would like to acknowledge the financial support provided by the Actuarial Research Club, particularly in respect of the participation of the first author in this project.

Appendix: AHP- A BRIEF DESCRIPTION

A1. Analytic Hierarchy Process (AHP) method is one of the widely used multi-criteria decision making (MCDM) methods and is based on the principle that in decision making, experience and knowledge of the decision makers is at least as valuable as the data, which describes the particular problem and on which the conclusions are normally based.

Applications of AHP are found in many areas and among these are bank strategic planning, benefit-cost evaluation, bond ratings models. We will give a brief description of the method, closely following chapter 2 of Kumar&Ganesh (1996). A detailed discussion of the theory of AHP is also presented by Saaty (1980).

A2. The AHP process can be thought of as consisting of four parts (Kumar and Ganesh, 1996).

First, the problem in investigation is decomposed hierarchically, starting from the overall objective's main characteristics (level n, where n is the number of levels) and each of the elements at a previous level is decomposed until the process gets to a point, where a reasonable evaluations of these 'basic' characteristics can be obtained.

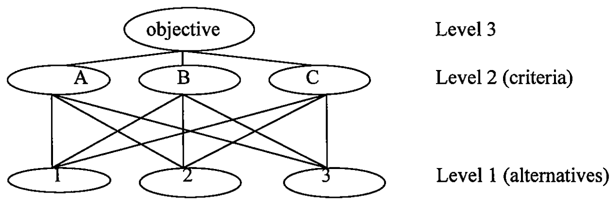


Fig. A1 AHP hierarchy.

Second, at each level, the pairs of elements are compared with respect to every element in the higher level, using Saaty's nine- point scale and a pairwise comparison matrix is defined. The matrix is reciprocal ($a_{ij}=1/a_{ji}$) and if it is consistent ($a_{ij}=a_{ik}a_{kj}$) then the computations in part 3 are straightforward. The matrix definition is a subjective operation and fully reflects the decision maker's knowledge and preferences. Some broad guidelines are given by Saaty's scale presented below.

Intensity of importance	Definition	Description
1	Equal importance	Two criteria contribute equally to the objective in the immediate higher level
3	Weak importance of one over another	Experience and judgement slightly favour one criterion over another
5	Essential or strong importance	Experience and judgement strongly favour one criterion over another
7	Very strong or demonstrated importance	A criterion is favoured very strongly; its dominance demonstrated in practice
9	Absolute importance	The evidence favouring one criterion over another is of the highest possible order of affirmation
2, 4, 6, 8	Intermediate values between adjacent scale values	When compromise is needed.
Reciprocals of the above judgements	If criterion A has one of the above judgements compared to criterion B then B has the reciprocal value when compared to A	A reasonable assumption

Table A1. Saaty's nine- point scale and its description.

Third, the eigenvector method is used for generating the weights/priority vectors at each level with respect to every element at the next higher level. The weights vector is the normalised principal eigenvector of a pairwise comparison matrix i.e.

w': $Aw'=\lambda w'$ and $w=w'/||w'||$ is the normalised principal eigenvector of A, where A is a reciprocal square matrix ($a_{ij}=1/a_{ji}$) and λ is its principal eigenvalue.

And *finally* all weight vectors are synthesised in order to obtain the final vector of weights. The weights vectors from a level are used to obtain the weights vectors at the next higher level. If W_i denotes the weight of alternative i , c_j is the weight of criterion j and w_{ij} is the weight of the alternative i with respect to criterion j, then

$$W_i = \sum c_j w_{ij} \quad .$$

A3. AHP method is not perfect and it has received a mixed reaction among researchers. The main drawback is the so called *rank reversal phenomenon*.This refers to the inability of the theory to preserve the ranking of the alternatives if an alternative is added or deleted. But the theory and the associated mathematics are relatively simple, the method is intuitively clear and easy to understand and it can be equally easy to deal with quantitative and qualitative data.

8. CONCLUSIONS AND FURTHER RESEARCH.

8.1 Actuarial science provides solutions to problems involving the financial implications of future uncertain events and makes wide use of mathematical and statistical techniques. The range of fields in which actuarial expertise is employed is both extensive and complex and some would argue, could be expanded (Ferguson, 1997). The arsenal of actuarial techniques needs a continuous update in order to keep a competitive edge in presentation of and solution finding in new and unfamiliar situations or to present in a new light the existing ones (Cook and Valentine, 1997). Fuzzy set theory provides such an opportunity.

8.2 Fuzzy sets are used to describe uncertain statements, where the uncertainty is due to the nature of the phenomenon, its perception by humans or arising from its complexity.
Since the concept of a fuzzy set was first introduced some thirty years ago, FST has undergone a rapid and continuous development and now it is common to find a fuzzy approach in linguistics, artificial intelligence, pattern analysis and classification, decision making and many other branches of contemporary science.

8.3 In this paper, we have presented a review of fuzzy concepts and techniques which have been used in an actuarial environment and we have presented new ideas, such as hybrid models, which might be of interest in future. The applications reviewed include areas such as financial mathematics, asset/ liability considerations, risk classification and underwriting in both life and non- life insurance.

8.4 FST is a new branch of modern mathematics and its applications in actuarial science are even more recent. Our view is that it provides a promising way of treating the uncertainty which is inherent in many actuarial applications. We believe that FST would be a useful addition to the modelling tools used by actuaries and that many of the potential applications lie in non-life insurance and the (so-called) 'wider' fields.

REFERENCES:

Babad, Y.M. and B. Berliner, 1994, The use of intervals of possibilities to measure and evaluate financial risks and uncertainty, 4th AFIR International Colloquium 1: 111 - 140.

Babad , Y.M. and B. Berliner, 1995, Reduction of uncertainty through actualization of intervals of possibilities and its applications to reinsurance, 25th TICA 2: 23 - 35.

Bellman, R. and L.A. Zadeh, 1970, Decision making in a fuzzy environment, Management Science 17: B141-B164.

Berliner, B. and N. Buehlmann, 1993, A generalization of the fuzzy zooming of cash flows, 3rd AFIR International Colloquium 2: 431 - 456.

Bezdek, J.C., 1981, Pattern recognition with fuzzy objective function algorithms, Plenum Press, New York, London.

Black, M., 1937, Vagueness: an exercise in logical analysis, Philosophy of Science, 427-455.

Buckley, J.J., 1987, The fuzzy mathematics of finance, Fuzzy sets and systems, 21:257-273.

Buehlmann, N.,B. Berliner, 1992, Fuzzy zooming of cash flows, 24th TICA, 6:437-453.

Chang, C.C. and P.Z. Wang., 1995, The matching of assets and liabilities with fuzzy mathematics, 25th TICA, 3:123 - 137.

Chen, J.J.G, and Z.He, 1997, Using analytic hierarchy process and fuzzy set theory to rate and rank the disability, Fuzzy Sets and Systems, 88:1-22.

Cook, N. and B. Valentine, 1997, Fractals: The physicists' wolf stalking the actuarial flock, Presented to Staple Inn Actuarial Society, London.

Cummins, J.D. and R.A. Derrig, 1993, Fuzzy trends in property - liability insurance claim costs, Journal of Risk and Insurance, 60: 429 - 465.

Cummins, J.D. and R.A. Derrig, 1997, Fuzzy financial pricing of property liability insurance, North-American Actuarial Journal, 1: 21-40.

Derrig, R.A. and K.M. Ostaszewski, 1994, Fuzzy techniques of pattern recognition in risk and claim classification, 4th AFIR International Colloquium, 1:141 - 171.

Derrig, R.A. and K.M. Ostaszewski, 1995, The fuzzy problem of hedging the tax liability of a property - liability insurance company, 5th AFIR International Colloquium, 1:17 - 42.

Derrig, R.A., 1997, personal communication.

DeWit, G.W., 1982, Underwriting and uncertainty, Insurance: Mathematics and Economics, 1:277-285.

Dubois, D.and H. Prade, 1980, Fuzzy sets and systems, Academic Press, San Diego.

Dumouchel, W.H., 1983, The 1992 Massachusetts automobile insurance classification scheme, The Statistician, 32: 69 - 81.

Ferguson, D.G.R., 1997, Presidential address, British Actuarial Journal, 3:1-26.

Geman, S., E. Bienenstock, and R. Doursat, 1992, Neural Networks and the Bias/Variance Dilemma, Neural Computation, 4:1-58.

Guo, L. and Z. Huang 1996, A possibilistic linear programming method for asset allocation, Journal of Actuarial Practice, 2:67-90.

Hellman, A., 1995, A fuzzy expert system for evaluation of municipalities - an application, 25th TICA, 1:159 - 187.

Horgby, P.J., R. Lohse, N.A. Sittaro, 1997, Fuzzy underwriting: an application of fuzzy logic to medical underwriting, Journal of Actuarial Practice, 5: 79-104.

Klir, G.J. and T.A.Folger, 1988, Fuzzy sets, uncertainty, and information, Prentice-Hall, New Jersey.

Kosko, B., 1990, Fuzziness vs. probability, International journal of general systems, 17:211-240.

Kramer, B., 1996, N.E.W.S.: a model for the evaluation of non-life insurance companies, Ph.D. thesis, University of Groningen, Groningen, the Netherlands.

Kumar, N.V. and L.S. Danesh, 1996, An empirical analysis of the use of the AHP for estimating membership values in a fuzzy set, Fuzzy sets and systems, 82:1- 16.

Lemaire, J., 1990, Fuzzy insurance, ASTIN Bulletin, 20:33-55.

Loimaranta, K., J. Jacobsson and H. Lonka, 1980, On the use of mixture models in clustering multivariate frequency data, TICA, 21:147- 161.

McCauley-Bell, P.and A. Badiru, 1996, Fuzzy modelling and Analytic Hierarchy Processing to Quantify Risk Levels Associated with Occupational Injuries, PartI and II, IEEE Transactions on Fuzzy Systems, 4:124-138.

Minsky, M., 1975, A framework for representing knowledge, In The psychology of computer vision, McGraw-Hill, New York.

Ostaszewski, K.M., 1993, An investigation into possible applications of fuzzy set methods in actuarial science, Society of Actuaries.

Saaty, T.L., 1980, The analytic hierarchy process, McGraw- Hill, New York.

Terano, T., K.Asai, M. Sugeno, 1987, Fuzzy systems theory and its applications, Harcourt-Brace, Boston, MA.

Waterman, D.A., 1986, A guide to expert systems, Addison-Wesley, Reading, MA.

Weisberg, H.I. and R.A. Derrig, 1992, Massachusetts Automobile Bodily Injury Tort Reform, Journal of Insurance Regulation, 10:384-440.

Werth, M., 1995, Preferred lives- a more complete method of risk assessment, Presented to Staple Inn Actuarial Society, London.

Yakoubov, Y.H., 1997, Grouping policyholder age in motor insurance-a fuzzy approach, Msc dissertation, City University, London.

Young, V.R., 1993, The application of fuzzy sets to group health underwriting, TSA, 45: 551 - 590.

Young, V.R., 1996, Insurance Rate changing: A Fuzzy Logic Approach, The Journal of Risk and Insurance, 63: 461 - 484.

Zadeh, L. A., 1965, Fuzzy sets, Information and Control, 8:338-353.

Zimmerman, H.J, 1991, Fuzzy set theory and its applications, Kluwer Academic Publishers, Boston/Dordrecht/London.

CITY University

ACTUARIAL RESEARCH CENTRE

DEPARTMENT OF ACTUARIAL SCIENCE AND STATISTICS

Actuarial Research Papers

1. Haberman S. Long Term Prognosis after Cerebral Infarction. 78 pages, 1984 (out of print).
ISBN 1 874 770 01 8

2. Haberman S. Long Term Prognosis after Subarachnoid Haemorrhage. 43 pages, 1984 (out of print).
ISBN 1 874 770 02 6

3. Haberman S. Age Variation in Incidence of Cerebrovascular Disease. 36 pages, 1985 (out of print).
ISBN 1 874 770 03 4

4. Kaye G.D. Taxation of Pension Funds. 11 pages, 1985.
ISBN 1 874 770 04 2

[Also: Haberman S. and Bloomfield D.S.F. Social Class Differences in Mortality Around 1981. SSRU Working Paper No. 55. 41 pages, 1987.]

5. Kaye G.D. Life Offices' Expenses Regression Analyses I, 98 pages, 1988. (Limited circulation).
ISBN 1 874 770 05 0

6. Kaye G.D. Life Assurance Taxation - Comments in Response to the Board of the Inland Revenue's Consultative Document Concerning the Taxation of Life Insurance. 11 pages. 1988.
ISBN 1 874 770 06 9

7. Verrall R.J. Bayesian Linear Models and the Claims Run-Off Triangle. 32 pages. 1988.
ISBN 1 874 770 07 7

8. Kaye G.D. Life Offices' Expenses Regression Analyses II. 12 pages. 1989. (Limited circulation).
ISBN 1 874 770 08 5

9. Bloomfield D.S.F. and Haberman S. The Assessment of Damages and the Measurement of Work time Lost to Sickness, Unemployment and Stoppages. 100 pages. 1989.
ISBN 1 874 770 09 3

10. Adams A.T. Pricing Property Investment Company Shares. 14 pages. 1989.
ISBN 1 874 770 10 7

11. Booth P.M. History of Exchange Controls in the EC with Particular Reference to the United Kingdom. 39 pages. 1989.
ISBN 1 874 770 11 5

12. Verrall R.J. Chain Ladder and Maximum Likelihood. 12 pages. 1989.
ISBN 1 874 770 12 3

13. Verrall R.J. On the Unbiased Estimation of Reserves from Loglinear Models. 13 pages, 1989.
ISBN 1 874 770 13 1

14. Adams A.T. An Analysis of Investment Trust Shares, 1990. 31 pages.
ISBN 1 874 770 14 X

15. Haberman S. Demographic Trends Towards 2000. 16 pages. 1989.
ISBN 1 874 770 15 8

16. Haberman S. British Life Insurance: The Role of the Actuary and 1992. 30 pages. 1990.
ISBN 1 874 770 16 6

17. Verrall R.J. Bootstrapping, Saddlepoint Approximations and the Probability of Ruin. 10 pages. July 1990.
ISBN 1 874 770 17 4

18. Booth P.M. Theoretical Aspects of the Taxation of Real Returns from Short Term Deposits. 26 pages. June 1990.
ISBN 1 874 770 18 2

19. Booth P.M. Practical Aspects of the Taxation of Real Interest. 38 pages, 1990.
ISBN 1 874 770 19 0

20. Puzey A.S. Actuarial Funding in Unit-linked Life Assurance - Derivation of the Actuarial Funding Factors. 19 pages. July 1990.
ISBN 1 874 770 20 4

21. Kaye G.D. Economies of Scale in the United Kingdom Insurance Industry. A report for the Association of British Insurers. Limited Circulation. October 1990.
ISBN 1 874 770 21 2

22. Adams A.T. & Venmore-Rowland P. Direct or Indirect Property Investment. 71 pages. September 1990.
ISBN 1 874 770 22 0

23. England P.D. Statistical Modelling of Excess Mortality. 124 pages. September 1990.
ISBN 1 874 770 23 9

24. Verrall R.J. & Li Z. Negative Incremental Claims: Chain Ladder and Linear Models. November 1990.
ISBN 1 874 770 24 7

25. Haberman S. State Pension Provision and Social Security in the United Kingdom. Presented at a Conference in Lisbon on "Social Security and Private Provision; What are the Prospects?". 40 pages. October 1990.
ISBN 1 874 770 25 5

26. Renshaw A.E. Graduation by Generalized Linear Modelling Techniques. 42 pages. February 1991.
ISBN 1 874 770 26 3

27. England P.D. Statistical Modelling of Excess Mortality No. 2. 170 pages. June 1991.
ISBN 1 874 770 27 1

28. Adams, A.T. The Market Pricing of Composite Insurance Shares. 37 pages. July 1991.
ISBN 1 874 770 28 X

29. Haberman, S. The Effect of Autoregressive Rates of Return on the Variability of Pension Contributions and Fund Levels. 58 pages. September 1991.
ISBN 1 874 770 29 8

30. Puzey, A.S. An Investigation into the Effect of Averaging Age Nearest Birthday at the Beginning of a Rate Year to Exact Age. October 1991. 33 pages.
ISBN 1 874 770 30 1

31. Chadburn, R.G. Bias in Select Mortality Investigations where Data are Subdivided by Age at Death. November 1991. 69 pages.
ISBN 1 874 770 31 X

32. England P.D. and Verrall R.J. Dynamic Estimation for Models of Excess Mortality. January 1992. 19 pages.
ISBN 1 874 770 32 8

33. Verrall R.J. A State Space Formulation of Whittaker-Henderson Graduation, with Extensions. January 1992. 13 pages.
ISBN 1 874 770 33 6

34. Verrall R.J. Graduation by Dynamic Regression Methods. January 1992. 37 pages.
ISBN 1 874 770 34 4

35. Gerrard R.G. and Haberman S. Stability of Pension Systems when Gains/Losses are Amortized and Rates of Return are Autoregressive. March 1992. 12 pages.
ISBN 1 874 770 35 2

36. Haberman S. HIV, AIDS and the Approximate Calculation of Life Insurance Functions, Annuities and Net Premiums. April 1992. 28 pages.
ISBN 1 874 770 36 0

37. Cooper D.R. Savings and Loans: An Historical Perspective. May 1992. 29 pages.
ISBN 1 874 770 37 9

38. Verrall R.J. Dynamic Bayes Estimation for AIDS Delay Tables. May 1992. 16 pages.
ISBN 1 874 770 38 7

39. Kaye G.D. Data Sources Relating to Life Assurance Expenses. May 1992. 39 pages. Presented to the Staple Inn Actuarial Society and Royal Statistical Society - February 1992.
ISBN 1 874 770 39 5

40. Renshaw A.E., and Haberman S. On the Graduation Associated with a Multiple State Model for Permanent Health Insurance. May 1992.
ISBN 1 874 770 40 9

41. England P.D. Statistical Modelling of Excess Mortality Number 3. June 1992. 163 pages.
ISBN 1 874 770 41 7

42. Bloomfield D.S.F. and Haberman S. Male Social Class Mortality Differences Around 1981: An Extension to Include Childhood Ages. 21 pages. June 1992.
ISBN 1 874 770 42 5

43. Berg M.P and Haberman S. Trend Analysis and Prediction Procedures for Time Nonhomogeneous Claim Processes. 33 pages. June 1992.
ISBN 1 874 770 43 3

44. Booth P.M. The Single Market for Insurance, Free Capital Movements and attempts to Fix Exchange Rates. October 1992. 28 pages.
ISBN 1 874 770 44 1

45. Verrall R.J. Chain Ladder with Varying Run-off Evolutions. February 1993. 15 pages.
ISBN 1 874 770 45 X

46. Gavin J., Haberman S. and Verrall R.J. Moving Weighted Average Graduation using Kernel Estimation. November 1992. 14 pages.
ISBN 1 874 770 46 8

47. Gavin J., Haberman S. and Verrall R.J. On the Choice of Bandwidth for Kernel Graduation. November 1992. 21 pages.
ISBN 1 874 770 47 6

48. S. Haberman. Pension Funding with Time Delays and the Optimal Spread Period. May 1993. 13 pages.
ISBN 1 874 770 48 4

49. S. Haberman. Stochastic Investment Returns and the Present Value of Future Contributions in a Defined Benefit Pension Scheme. May 1993. 22 pages.
ISBN 1 874 770 49 2

50. A. Zimbidis and S. Haberman. Delay, Feedback and Variability of Pension Contributions and Fund Levels. May 1993. 19 pages.
ISBN 1 874 770 50 6

51. S. Haberman. Pension Funding: The Effect of Changing The Frequency Valuations. June 1993. 19 pages.
ISBN 1 874 770 51 4

52. S Haberman. HIV, AIDS Markov Chains and PHI. June 1993. 15 pages.
ISBN 1 874 770 52 2

53. S Haberman. A Consideration of Pension Credit and Termination Insurance. June 1993. 22 pages.
ISBN 1 874 770 53 0

54. M Z Khorasanee. Survey of Actuarial Practice in the Funding of UK Defined Benefit Pension Schemes. July 1993. 19 pages.
ISBN 1 874 770 54 9

55. P M Booth, R G Chadburn and A S K Ong. A Utility Maximisation Approach to Asset Allocation. September 1993. 40 pages.
ISBN 1 874 770 55 7

56. R G Chadburn. Bias in Select Mortality Investigations. August 1993. 62 pages.
ISBN 1 874 770 56 5

57. M Z Khorasanee. A Comparison of Pension Funding Strategies by means of Simulations for a Model Scheme. August 1993. 43 pages.
ISBN 1 874 770 57 3

58. A E Renshaw, P Hatzopolous and S Haberman. Recent Mortality Trends in Male Assured Lives. June 1993. 23 pages.
ISBN 1 874 770 58 1

59. E Pitacco. Disability Risk Models: Towards A Unifying Approach. September 1993. 33 pages.
ISBN 1 874 770 59 X

60. M Boskov and R J Verrall. Premium Rating by Geographic Area Using Spatial Models. September 1993. 14 pages.
ISBN 1 874 770 60 3

61. R G Chadburn. Managing Mutual Life Offices: The Use of an Expense Ratio in New Business Decision Making and Expense Control. October 1993. 21 pages.
ISBN 1 874 770 61 1

62. Haberman S. Pension Funding Modelling and Stochastic Investment Returns. 56 pages. March 1994.
ISBN 1 874 770 62 X

63. Renshaw A E and Verrall R J. The Stochastic Model Underlying the Chain-Ladder Technique. 25 pages. April 1994.
ISBN 1 874 770 63 8

64. Haberman S and Sung J-H. Dynamic Approaches to Pension Funding. 22 pages. April 1994.
ISBN 1 874 770 64 6

65. Renshaw A.E. On the Second Moment Properties and the Implementation of Certain GLIM Based Stochastic Claims Reserving Models. 36 pages. September 1994.
ISBN 1 874 770 65 4

66. Booth P.M., J.N. Allan, and J.W. Jang. An Evaluation of the UK Life Insurance Mismatching Test. September 1994.
ISBN 1 874 770 66 2

67. Booth P.M. and Stroinski K. Insurance and Investment Markets in Poland. September 1994. 35 pages.
ISBN 1 874 770 67 0

68. Ong A. A Stochastic Model for Treasury-Bills: An Extension to Wilkie's Model. September 1994. 12 pages.
ISBN 1 874 770 68 9

69. Bloomfield D.S.F. Moving on from Undergraduate Exams to Professional Exams: Actuaries. November 1994. 22 pages.
ISBN 1 874 770 69 7

70. Huber P. A Review of Wilkie's Stochastic Investment Model. January 1995. 22 pages.
ISBN 1 874 770 70 0

71. Renshaw A.E. On the Graduation of 'Amounts'. January 1995. 24 pages.
ISBN 1 874 770 71 9

72. Renshaw A.E. Claims Reserving by Joint Modelling. December 1994. 26 pages.
ISBN 1 874 770 72 7

73. Renshaw A.E. Graduation and Generalised Linear Models: An Overview. February 1995. 40 pages.
ISBN 1 874 770 73 5

74. Khorasanee M.Z. Simulation of Investment Returns for a Money Purchase Fund. June 1995. 20 pages.
ISBN 1 874 770 74 3

75. Owadally M.I. and Haberman S. Finite-time Pension Fund Dynamics with Random Rates of Return. June 1995. 28 pages.
ISBN 1 874 770 75 1

76. Owadally M.I. and Haberman S. Stochastic Investment Modelling and Optimal Funding Strategies. June 1995. 25 pages.
ISBN 1 874 770 76 X

77. Khorasanee M.Z Applying the Defined Benefit Principle to a Defined Contribution Scheme. August 1995. 30 pages.
ISBN 1 874 770 77 8

78. Sebastiani P. and Settimi R. Experimental Design for Non-Linear Problems. September 1995. 13 pages.
ISBN 1 874 770 78 6

79. Verrall R.J. Whittaker Graduation and Parametric State Space Models. November 1995. 23 pages.
ISBN 1 874 770 79 4

80. Verrall R.J. Claims Reserving and Generalised Additive Models. November 1995. 17 pages.
ISBN 1 874 770 80 8

81. Nelder J.A. and Verrall R.J. Credibility Theory and Generalized Linear Models. November 1995. 15 pages.
ISBN 1 874 770 81 6

82. Renshaw A.E., Haberman S. and Hatzopoulos P. On The Duality of Assumptions Underpinning The Construction of Life Tables. December 1995. 17 Pages.
ISBN 1 874 770 82 4

83. Chadburn R.G. Use of a Parametric Risk Measure in Assessing Risk Based Capital and Insolvency Constraints for With Profits Life Insurance. March 1996. 17 Pages.
ISBN 1 874 770 84 0

84. Haberman S. Landmarks in the History of Actuarial Science (up to 1919). March 1996. 62 Pages.
ISBN 1 874 770 85 9

85. Renshaw A.E. and Haberman S. Dual Modelling and Select Mortality. March 1996. 30 Pages.
ISBN 1 874 770 88 3

86. Booth P.M. Long-Term Care for the Elderly: A Review of Policy Options. April 1996. 45 Pages.
ISBN 1 874 770 89 1

87. Huber P.P. A Note on the Jump-Equilibrium Model. April 1996. 17 Pages.
ISBN 1 874 770 90 5

88. Haberman S and Wong L.Y.P. Moving Average Rates of Return and the Variability of Pension Contributions and Fund Levels for a Defined Benefit Pension Scheme. May 1996. 51 Pages.
ISBN 1 874 770 91 3

89. Cooper D.R. Providing Pensions for Employees with Varied Working Lives. June 1996. 25 Pages.
ISBN 1 874 770 93 X

90. Khorasanee M.Z. Annuity Choices for Pensioners. August 1996. 25 Pages.
ISBN 1 874 770 94 8

91. Verrall R.J. A Unified Framework for Graduation. November 1996. 25 Pages.
ISBN 1 874 770 99 9

92. Haberman S. and Renshaw A.E. A Different Perspective on UK Assured Lives Select Mortality. November 1996. 61 Pages.
ISBN 1 874 770 00 X

93. Booth P.M. The Analysis of Actuarial Investment Risk. March 1997. 43 Pages.
ISBN 1 901615 03 0

94. Booth P.M., Chadburn R.G. and Ong A.S.K. Utility-Maximisation and the Control of Solvency for Life Insurance Funds. April 1997. 39 Pages.
ISBN 1 901615 04 9

95. Chadburn R.G. The Use of Capital, Bonus Policy and Investment Policy in the Control of Solvency for With-Profits Life Insurance Companies in the UK. April 1997. 29 Pages.
ISBN 1 901615 05 7

96. Renshaw A.E. and Haberman S. A Simple Graphical Method for the Comparison of Two Mortality Experiences. April 1997. 32 Pages.
ISBN 1 901615 06 5

97. Wong C.F.W. and Haberman S. A Short Note on Arma (1, 1) Investment Rates of Return and Pension Funding. April 1997. 14 Pages.
ISBN 1 901615 07 3

98. Puzey A S. A General Theory of Mortality Rate Estimators. June 1997. 26 Pages.
ISBN 1 901615 08 1

99. Puzey A S. On the Bias of the Conventional Actuarial Estimator of q_x. June 1997. 14 Pages.
ISBN 1 901615 09 X

100. Walsh D. and Booth P.M. Actuarial Techniques in Pricing for Risk in Bank Lending. June 1997. 55 Pages.
ISBN 1 901615 12 X

101. Haberman S. and Walsh D. Analysis of Trends in PHI Claim Inception Data. July 1997. 51 Pages.
ISBN 1 901615 16 2

102. Haberman S. and Smith D. Stochastic Investment Modelling and Pension Funding: A Simulation Based Analysis. November 1997. 91 Pages.
ISBN 1 901615 19 7

103. Rickayzen B.D. A Sensitivity Analysis of the Parameters used in a PHI Multiple State Model. December 1997. 18 Pages.
ISBN 1 901615 20 0

104. Verrall R.J. and Yakoubov Y.H. A Fuzzy Approach to Grouping by Policyholder Age in General Insurance. January 1998. 18 Pages.
ISBN 1 901615 22 7

105. Yakoubov Y.H. and Haberman S. Review of Actuarial Applications of Fuzzy Set Theory. February 1998. 88 Pages.
ISBN 1 901615 23 5

Statistical Research Papers

1. Sebastiani P. Some Results on the Derivatives of Matrix Functions. December 1995. 17 Pages.
ISBN 1 874 770 83 2

2. Dawid A.P. and Sebastiani P. Coherent Criteria for Optimal Experimental Design. March 1996. 35 Pages.
ISBN 1 874 770 86 7

3. Sebastiani P. and Wynn H.P. Maximum Entropy Sampling and Optimal Bayesian Experimental Design. March 1996. 22 Pages.
ISBN 1 874 770 87 5

4. Sebastiani P. and Settimi R. A Note on D-optimal Designs for a Logistic Regression Model. May 1996. 12 Pages.
ISBN 1 874 770 92 1

5. Sebastiani P. and Settimi R. First-order Optimal Designs for Non Linear Models. August 1996. 28 Pages.
ISBN 1 874 770 95 6

6. Newby M. A Business Process Approach to Maintenance: Measurement, Decision and Control. September 1996. 12 Pages.
ISBN 1 874 770 96 4

7. Newby M. Moments and Generating Functions for the Absorption Distribution and its Negative Binomial Analogue. September 1996. 16 Pages.
ISBN 1 874 770 97 2

8. Cowell R.G. Mixture Reduction via Predictive Scores. November 1996. 17 Pages.
ISBN 1 874 770 98 0

9. Sebastiani P. and Ramoni M. Robust Parameter Learning in Bayesian Networks with Missing Data. March 1997. 9 Pages.
ISBN 1 901615 00 6

10. Newby M.J. and Coolen F.P.A. Guidelines for Corrective Replacement Based on Low Stochastic Structure Assumptions. March 1997. 9 Pages.
ISBN 1 901615 01 4.

11. Newby M.J. Approximations for the Absorption Distribution and its Negative Binomial Analogue. March 1997. 6 Pages.
ISBN 1 901615 02 2

12. Ramoni M. and Sebastiani P. The Use of Exogenous Knowledge to Learn Bayesian Networks from Incomplete Databases. June 1997. 11 Pages.
ISBN 1 901615 10 3

13. Ramoni M. and Sebastiani P. Learning Bayesian Networks from Incomplete Databases. June 1997. 14 Pages.
ISBN 1 901615 11 1

14. Sebastiani P. and Wynn H.P. Risk Based Optimal Designs. June 1997. 10 Pages.
ISBN 1 901615 13 8

15. Cowell R. Sampling without Replacement in Junction Trees. June 1997. 10 Pages.
ISBN 1 901615 14 6

16. Dagg R.A. and Newby M.J. Optimal Overhaul Intervals with Imperfect Inspection and Repair. July 1997. 11 Pages.
ISBN 1 901615 15 4

17. Sebastiani P. and Wynn H.P. Bayesian Experimental Design and Shannon Information. October 1997. 11 Pages.
ISBN 1 901615 17 0

18. Wolstenholme L.C. A Characterisation of Phase Type Distributions. November 1997. 11 Pages.
ISBN 1 901615 18 9

19. Wolstenholme L.C. A Comparison of Models for Probability of Detection (POD) Curves. December 1997. 23 Pages.
ISBN 1 901615 21 9

ISBN 1 901615 23 5

Department of Actuarial Science and Statistics

Actuarial Research Club

The support of the corporate members

Commercial Union
Coopers & Lybrand
Government Actuary's Department
Guardian Insurance
Hymans Robertson
KPMG
Munich Reinsurance
Swiss Reinsurance

is gratefully acknowledged.